



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 118620

TO: Manjunath N Rao
Location: REM-3B81&3C70
Art Unit: 1652
Wednesday, April 07, 2004

Case Serial Number: 09/246451

From: Noble Jarrell
Location: Biotech-Chem Library
Rem 1B71
Phone: 272-2556

Noble.jarrell@uspto.gov

Search Notes

STIC-Biotech/ChemLib

118820

From: Rao, Manjunath N.
Sent: Tuesday, April 06, 2004 10:54 AM
To: STIC-Biotech/ChemLib
Subject: Sequence search request for 09/246,451

From: Manjunath N. Rao
Art Unit 1652, Room 3B-81
Phone: 571-272-0939

Date: 4-6-04

Please search the following as soon as possible for application with serial number **09/246,451**

SEQ ID NOs. 2, 11, 12, 13 and 17 against all protein databases including issued patents database, published applications and pending application database and provide a print of all results.

If you have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao
Art Unit 1652
Room 3B-81, Remsen Bldg.
Phone: 272-0939

Manjunath N. Rao
Art Unit 1652, Room 3B81,
Mail Box in Room 3070,
Remsen Building, USPTO
400, Dulany St.
Alexandria, VA.
Phone: 571-272-0939

Searcher: Garrett/Schreiber
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 4/7/04
Searcher Prep/Review: 5
Clerical: _____
Online time: 5

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: 5
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: CompuGen
WWW/Internet: _____
Other (specify): _____

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 18:53:57 ; Search time 34.974 Seconds
(without alignments)
3108.883 Million cell updates/sec

Title: US-09-246-451a-2

Perfect score: 2180

Sequence: 1 TTTTOSNANLAPLPHVPE.....IVSGVQALVWDPATTKAV 414

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 262633353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0

Maximum DB seq length: 2006000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US05_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US04_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US03_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US02_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US01_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US00_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US05_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US04_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US03_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US02_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US01_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US00_PUBCOMB.pep.*

2red. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2180	100.0	414	15	US-10-453-104-2
2	2176	99.8	414	15	US-10-453-104-11
3	2169	99.5	414	15	US-10-453-104-12
4	2165	99.3	414	15	US-10-453-104-13
5	377.5	17.2	404	14	US-10-214-446-50
6	374.5	17.2	404	14	US-10-214-446-40
7	372	17.1	399	14	US-10-314-657-37
8	367	16.8	416	14	US-10-156-761-14997
9	363.5	16.7	416	9	US-09-861-289-39
10	363.5	16.7	416	9	US-09-860-846-39
11	363.5	16.7	416	10	US-09-988-384B-39
12	363.5	16.7	416	10	US-09-836-821-39
13	363.5	16.7	416	10	US-09-793-708-18
14	363.5	16.7	416	14	US-10-201-365-13
15	363.5	16.7	416	14	US-10-163-539-18

16 363.5 16.7 416 14 US-10-271-889-39 Sequence 39, Appl
17 359.5 16.5 411 14 US-10-156-761-8376 Sequence 39, Appl
18 356 16.3 403 14 US-10-156-761-9525 Sequence 39, Appl
19 355 16.3 403 14 US-10-214-446-38 Sequence 39, Appl
20 354 16.2 425 14 US-10-214-446-20 Sequence 39, Appl
21 352.5 16.2 457 14 US-10-156-761-11073 Sequence 39, Appl
22 348 16.3 418 12 US-10-389-647-559 Sequence 39, Appl
23 347.5 15.9 399 14 US-10-156-761-9914 Sequence 39, Appl
24 339.5 15.6 388 14 US-10-156-761-13776 Sequence 39, Appl
25 336.5 15.4 408 14 US-10-214-446-4 Sequence 39, Appl
26 333 15.3 409 15 US-10-458-201-12 Sequence 39, Appl
27 330 15.1 392 14 US-10-214-446-32 Sequence 39, Appl
28 328.5 15.1 406 14 US-10-214-446-16 Sequence 39, Appl
29 321 14.7 404 14 US-10-214-446-2 Sequence 39, Appl
30 320 14.7 401 14 US-10-156-761-8710 Sequence 39, Appl
31 318 14.6 428 14 US-10-201-213-6 Sequence 39, Appl
32 317.5 14.6 404 14 US-10-156-761-14659 Sequence 39, Appl
33 316.5 14.5 396 14 US-10-214-446-14 Sequence 39, Appl
34 315 14.4 412 14 US-10-214-446-36 Sequence 39, Appl
35 315 14.4 421 14 US-10-156-761-9703 Sequence 39, Appl
36 314 14.4 418 14 US-10-214-446-22 Sequence 39, Appl
37 307 14.1 421 14 US-10-214-446-18 Sequence 39, Appl
38 304.5 14.0 400 14 US-10-314-657-62 Sequence 39, Appl
39 304.5 14.0 404 14 US-10-156-761-10431 Sequence 39, Appl
40 300.5 13.8 415 14 US-10-214-446-46 Sequence 39, Appl
41 299 13.7 475 14 US-10-145-415-22 Sequence 39, Appl
42 298.5 13.7 402 14 US-10-205-032-8 Sequence 39, Appl
43 298 13.7 470 14 US-10-145-415-6 Sequence 39, Appl
44 296 13.6 429 14 US-10-145-415-14 Sequence 39, Appl
45 295 13.5 399 14 US-10-156-761-7959 Sequence 39, Appl

ALIGNMENTS

RESULT 1
US-10-453-104-2
; Sequence 2, Application US/10453104
; Publication No. US20030207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Hyun Joo
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
; FILE REFERENCE: 4058/1827-US3
; CURRENT APPLICATION NUMBER: US/10/453,104
; CURRENT FILING DATE: 2003-06-02
; PRIOR FILING DATE: 2000-09-13
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 09/246,451
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/086,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/106,834
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 414
; TYPE: PRT
; ORGANISM: P. Putida
US-10-453-104-2

Query Match 100.0%; Score 2180; DB 15; Length 414;
Best Local Similarity 100.0%; Pred. No. 1.3e-211;
Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TTTTOSNANLAPLPHVPEHLVDFEDMYNPSLISAGVQAWLQESVDFLWPTKNG 60

```

Db      1  TTETIQSNANLAPLPHPVPEHLVDFDMYNPNLSAGVQEAVALQESNVDPDLVWTRCNG 60
QY      61  GHWIATRGQILREAYEDYRHFSSCEPFPREAGEAYDFTPTSDPPEQRFALANQVVG 120
Db      61  GHWIATRGQILREAYEDYRHFSSCEPFPREAGEAYDFTPTSDPPEQRFALANQVVG 120
QY      121  MPVVDKLENRIQELACSLIESLRPOGCNFTEDYAEPPFIRIFMLLAGLPEEDIPHLKYL 180
Db      121  MPVVDKLENRIQELACSLIESLRPOGCNFTEDYAEPPFIRIFMLLAGLPEEDIPHLKYL 180
QY      181  TDQWTRPDGSMFAEAKAALYDYLPIIEQRQKPGTDAISIVANGVNGRPITSDAKR 240
Db      181  TDQWTRPDGSMFAEAKAALYDYLPIIEQRQKPGTDAISIVANGVNGRPITSDAKR 240
QY      241  MCGLLLVGGDLTVNVLFSFMEFLAKSPEHRQELIERPERIPAAACELLRFRSLVADGRI 300
Db      241  MCGLLLVGGDLTVNVLFSFMEFLAKSPEHRQELIERPERIPAAACELLRFRSLVADGRI 300
QY      301  LSTDYEFHGVQKKGDDQILLPQMLSGLDERENACPMHVDFSRQKVSHTTFGHSHLCLGQ 360
Db      301  LSTDYEFHGVQKKGDDQILLPQMLSGLDERENACPMHVDFSRQKVSHTTFGHSHLCLGQ 360
QY      361  HLARREIIVTLKWLTRIPDFSIAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414
Db      361  HLARREIIVTLKWLTRIPDFSIAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414

```

RESULT 2

```

US-10-453-104-11
: Sequence 11, Application US/10453104
: Publication No. US20030207345A1
: GENERAL INFORMATION:
: APPLICANT: California Institute of Technology;
: APPLICANT: Frances H. Arnold
: APPLICANT: Hyun Joo
: TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
: FILE REFERENCE: 4058/1827-US3
: CURRENT APPLICATION NUMBER: US/10/453,104
: CURRENT FILING DATE: 2003-06-02
: PRIOR APPLICATION NUMBER: US 09/661,093
: PRIOR FILING DATE: 2000-09-13
: PRIOR APPLICATION NUMBER: US 09/246,451
: PRIOR FILING DATE: 1999-02-09
: PRIOR APPLICATION NUMBER: US 60/094,403
: PRIOR FILING DATE: 1998-07-28
: PRIOR APPLICATION NUMBER: US 60/106,840
: PRIOR FILING DATE: 1998-11-03
: PRIOR APPLICATION NUMBER: US 60/086,206
: PRIOR FILING DATE: 1998-05-21
: PRIOR APPLICATION NUMBER: US 60/106,834
: PRIOR FILING DATE: 1998-11-03
: SOFTWARE: FastSeq for Windows Version 3.0
: NUMBER OF SEQ ID NOS: 19
: SEQ ID NO 11
: LENGTH: 414
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Mutant M7-4H
US-10-453-104-11

```

```

Query Match      99.8%; Score 2176; DB 15; Length 414;
Best Local Similarity 99.8%; Pred. No. 3.3e-211;
Matches 413; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1  TTETIQSNANLAPLPHPVPEHLVDFDMYNPNLSAGVQEAVALQESNVDPDLVWTRCNG 60
Db      1  TTETIQSNANLAPLPHPVPEHLVDFDMYNPNLSAGVQEAVALQESNVDPDLVWTRCNG 60
QY      61  GHWIATRGQILREAYEDYRHFSSCEPFPREAGEAYDFTPTSDPPEQRFALANQVVG 120
Db      61  GHWIATRGQILREAYEDYRHFSSCEPFPREAGEAYDFTPTSDPPEQRFALANQVVG 120

```

```

QY      121  MPVVDKLENRIQELACSLIESLRPOGCNFTEDYAEPPFIRIFMLLAGLPEEDIPHLKYL 180
Db      121  MPVVDKLENRIQELACSLIESLRPOGCNFTEDYAEPPFIRIFMLLAGLPEEDIPHLKYL 180
QY      181  TDQWTRPDGSMFAEAKAALYDYLPIIEQRQKPGTDAISIVANGVNGRPITSDAKR 240
Db      181  TDQWTRPDGSMFAEAKAALYDYLPIIEQRQKPGTDAISIVANGVNGRPITSDAKR 240
QY      241  MCGLLLVGGDLTVNVLFSFMEFLAKSPEHRQELIERPERIPAAACELLRFRSLVADGRI 300
Db      241  MCGLLLVGGDLTVNVLFSFMEFLAKSPEHRQELIERPERIPAAACELLRFRSLVADGRI 300
QY      301  LSTDYEFHGVQKKGDDQILLPQMLSGLDERENACPMHVDFSRQKVSHTTFGHSHLCLGQ 360
Db      301  LSTDYEFHGVQKKGDDQILLPQMLSGLDERENACPMHVDFSRQKVSHTTFGHSHLCLGQ 360
QY      361  HLARREIIVTLKWLTRIPDFSIAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414
Db      361  HLARREIIVTLKWLTRIPDFSIAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414

```

RESULT 3

```

US-10-453-104-12
: Sequence 12, Application US/10453104
: Publication No. US20030207345A1
: GENERAL INFORMATION:
: APPLICANT: California Institute of Technology;
: APPLICANT: Frances H. Arnold
: APPLICANT: Hyun Joo
: TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
: FILE REFERENCE: 4058/1827-US3
: CURRENT APPLICATION NUMBER: US/10/453,104
: CURRENT FILING DATE: 2003-06-02
: PRIOR APPLICATION NUMBER: US 09/661,093
: PRIOR FILING DATE: 2000-09-13
: PRIOR APPLICATION NUMBER: US 09/246,451
: PRIOR FILING DATE: 1999-02-09
: PRIOR APPLICATION NUMBER: US 60/094,403
: PRIOR FILING DATE: 1998-07-28
: PRIOR APPLICATION NUMBER: US 60/106,840
: PRIOR FILING DATE: 1998-11-03
: PRIOR APPLICATION NUMBER: US 60/086,206
: PRIOR FILING DATE: 1998-05-21
: PRIOR APPLICATION NUMBER: US 60/106,834
: PRIOR FILING DATE: 1998-11-03
: SOFTWARE: FastSeq for Windows Version 3.0
: NUMBER OF SEQ ID NOS: 19
: SEQ ID NO 12
: LENGTH: 414
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Mutant M7-6H
US-10-453-104-12

```

```

Query Match      99.5%; Score 2169; DB 15; Length 414;
Best Local Similarity 99.5%; Pred. No. 1.7e-210;
Matches 412; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1  TTETIQSNANLAPLPHPVPEHLVDFDMYNPNLSAGVQEAVALQESNVDPDLVWTRCNG 60
Db      1  TTETIQSNANLAPLPHPVPEHLVDFDMYNPNLSAGVQEAVALQESNVDPDLVWTRCNG 60
QY      61  GHWIATRGQILREAYEDYRHFSSCEPFPREAGEAYDFTPTSDPPEQRFALANQVVG 120
Db      61  GHWIATRGQILREAYEDYRHFSSCEPFPREAGEAYDFTPTSDPPEQRFALANQVVG 120
QY      121  MPVVDKLENRIQELACSLIESLRPOGCNFTEDYAEPPFIRIFMLLAGLPEEDIPHLKYL 180
Db      121  MPVVDKLENRIQELACSLIESLRPOGCNFTEDYAEPPFIRIFMLLAGLPEEDIPHLKYL 180
QY      181  TDQWTRPDGSMFAEAKAALYDYLPIIEQRQKPGTDAISIVANGVNGRPITSDAKR 240

```


Db 181 TDQMRPDGSMTFEAXKALDYLLPIIEQRKQKPGTDAISIVANGVNGRPITSDAKR 240
 QY 241 MCGLLLVGGLDTVWVNFSLFSMEFJAKSPHROELIERPERIPAAEELLRRSLVADGRI 300
 Db 241 MCGLLLVGGLDTVWVNFSLFSMEFJAKSPHROELIERPERIPAAEELLRRSLVADGRI 300
 QY 301 LTSDFEFGVQLKKGQDILLPQMLSGLDERENACPMHVDPSRQKVSHTTEGHSGLCLGQ 360
 Db 301 LTSDFEFGVQLKKGQDILLPQMLSGLDERENACPMHVDPSRQKVSHTTEGHSGLCLGQ 360
 QY 361 HJARREIIVTLKEMWTRIPDPSIAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414
 Db 361 HJARREIIVTLKEMWTRIPDPSIAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414

RESULT 4

US-10-453-104-13
 ; Sequence 13, Application US/10453104
 ; Publication No. US20030207345A1
 ; GENERAL INFORMATION:
 ; APPLICANT: California Institute of Technology;
 ; APPLICANT: Frances H. Arnold
 ; APPLICANT: Hyun Joo
 ; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
 ; FILE REFERENCE: 4058/1E827-US3
 ; CURRENT APPLICATION NUMBER: US/10/453,104
 ; CURRENT FILING DATE: 2003-06-02
 ; PRIOR APPLICATION NUMBER: US 09/651,093
 ; PRIOR FILING DATE: 2000-09-13
 ; PRIOR APPLICATION NUMBER: US 09/246,451
 ; PRIOR FILING DATE: 1999-02-09
 ; PRIOR APPLICATION NUMBER: US 60/094,403
 ; PRIOR FILING DATE: 1998-07-28
 ; PRIOR APPLICATION NUMBER: US 60/106,840
 ; PRIOR FILING DATE: 1998-11-03
 ; PRIOR APPLICATION NUMBER: US 60/086,206
 ; PRIOR FILING DATE: 1998-05-21
 ; PRIOR APPLICATION NUMBER: US 60/106,834
 ; PRIOR FILING DATE: 1998-11-03
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 13
 ; LENGTH: 414
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Mutant M7-8H
 US-10-453-104-13

Query Match 99.3%; Score 2165; DB 15; Length 414;
 Best Local Similarity 99.5%; Pred. No. 4.3e-210;
 Matches 412; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 TTETIGSNAMLAPLPHVPHLVPDPMYNSLACVQAWVLOESNVPDLVWTRCNG 60
 Db 1 TTETIGSNAMLAPLPHVPHLVPDPMYNSLACVQAWVLOESNVPDLVWTRCNG 60
 QY 61 GHWIATGQLIRAEYEDYRHSFSECCFPIPEAGEAYDFIPTMDPPQORFRALANOVVG 120
 Db 61 GHWIATGQLIRAEYEDYRHSFSECCFPIPEAGEAYDFIPTMDPPQORFRALANOVVG 120
 QY 121 MPVVDKLENRIQELASLIESLRPQGCNFTDYAEFPPIRIEMLLAGLPEEDIPHLKYL 180
 Db 121 MPVVDKLENRIQELASLIESLRPQGCNFTDYAEFPPIRIEMLLAGLPEEDIPHLKYL 180
 QY 181 TDQMRPDGSMTPAEAKALDYLLPIIEQRKQKPGTDAISIVANGVNGRPITSDAKR 240
 Db 181 TDQMRPDGSMTPAEAKALDYLLPIIEQRKQKPGTDAISIVANGVNGRPITSDAKR 240
 QY 241 MCGLLLVGGLDTVWVNFSLFSMEFJAKSPHROELIERPERIPAAEELLRRSLVADGRI 300
 Db 241 MCGLLLVGGLDTVWVNFSLFSMEFJAKSPHROELIERPERIPAAEELLRRSLVADGRI 300

QY 301 LTSDFEFGVQLKKGQDILLPQMLSGLDERENACPMHVDPSRQKVSHTTEGHSGLCLGQ 360
 Db 301 LTSDFEFGVQLKKGQDILLPQMLSGLDERENACPMHVDPSRQKVSHTTEGHSGLCLGQ 360
 QY 361 HJARREIIVTLKEMWTRIPDPSIAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414
 Db 361 HJARREIIVTLKEMWTRIPDPSIAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414

RESULT 5

US-10-214-446-50
 ; Sequence 50, Application US/10214446
 ; Publication No. US20030180742A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Weiner, David
 ; APPLICANT: Burk, Mark J.
 ; APPLICANT: Hitchman, Tim
 ; APPLICANT: Pujol, Catherine
 ; APPLICANT: Richardson, Toby
 ; APPLICANT: Short, Jay M.
 ; TITLE OF INVENTION: P450 ENZYMES, NUCLEIC ACIDS ENCODING
 ; TITLE OF INVENTION: THEM AND METHODS OF MAKING AND USING THEM
 ; FILE REFERENCE: 09010-500001
 ; CURRENT APPLICATION NUMBER: US/10/214,446
 ; CURRENT FILING DATE: 2002-08-05
 ; PRIOR APPLICATION NUMBER: US 60/309,497
 ; PRIOR FILING DATE: 2001-08-03
 ; NUMBER OF SEQ ID NOS: 59
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 50
 ; LENGTH: 404
 ; TYPE: PRT
 ; ORGANISM: Bacterial
 US-10-214-446-50

Query Match 17.3%; Score 377.5; DB 14; Length 404;
 Best Local Similarity 27.6%; Pred. No. 4e-29;
 Matches 97; Conservative 64; Mismatches 157; Indels 33; Gaps 5;

QY 60 GHWIATGQLIRAEYEDYRHSFSECCFPIPEAGEAYDFIPTSM----- 103
 Db 32 GRTWFLPHADIRTAALDERFSAS-----RKAGFVNFQFAEVPPEAFNEAISRWIV 85
 QY 104 --DPPEQORFRALANQVGMVVDKLENRIQELASLIESLRPQGCNFTDYAEFPPIR 161
 Db 86 LHDQPEHRQLRQLNCCQGFTRRLITTEPKIQKVCDDLIDAFVKRGSTEFMTEYAHFFPAK 145
 QY 162 IFMLLAGLPEEDIPHLKYLTDQMRPDGSM-----TFRAEKE---ALDYLLPIIEQBRQ 213
 Db 146 VIAEMLVNPEYDPAFVVMSEDLNFAAGSLRPTLEMFRAAQDGLLMMDYFARLLPERE 205
 QY 214 KPQTDALSIIVANGVNGRPITSDAKEMCGLLLVGGLDTVWVNFSLFSMEFLAKSPHROE 273
 Db 206 NPGDDLVSLLLSAESEGEWMTAEQVLANCTQIIVAGHETTUNLVANGVELLIRPEQAL 265
 QY 274 LTERPERIPAAEBELLRRFS-LVADGRILFSDYDFHGVQLKKGQDILLPQMLSGLDEREN 332
 Db 266 LESRPELMPSAVREIMEFSPLOQIRVAREDFEFGCAEVREGDGLVLMLSANRDEPAF 325
 QY 333 ACPMHVDPSRQKVSHTTEGHSGLCLGQHLARREIIVTLKEMWTRIPDPSI 383
 Db 326 DDPTDPTLRNPTGHIAFGHGVHVCVGAALAELEGGVFSRFLDLRLPGLLEL 376

RESULT 6

US-10-214-446-40
 ; Sequence 40, Application US/10214446
 ; Publication No. US20030180742A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Weiner, David
 ; APPLICANT: Burk, Mark J.
 ; APPLICANT: Hitchman, Tim
 ; APPLICANT: Pujol, Catherine

QY	174	IPHLKYLTDQWTRPDGSMTFAPAEKALDYLIPIITEQRRQKPGTDAISIVANGVQNGRPI	233
Db	181	HDFFESQSRLRFGIAEVDQARAQLDXYIYALIDRKKKEPGDGLIDDLIQEQLNGV	240
QY	234	TSDAKAMCGELLVGGGLDVTVNFSLFSMEFLAKSPHQELIERPERIPAACEEILRRFS	293
Db	241	DRAEVLSLATLLTAGHETTANM::SLGTFTLRHBEQLAEARPEGLMPAAVEELL::RFL	299
QY	294	LWADG--RILTSDEPHGVOLKKGQDILLPQWLSGLDERENACPHVDFPSRQKVSHITFG	351
Db	300	SIADGLLRVATEDIEVAGCTTIRADEGVVFATSVINRDAAFGAEPDADLWHSARHIVAFG	359
QY	352	HGSHLCISQHLARGLIIVLKEWTRIPDFSI-APGAQIQHSG-IVSGVQALPLW	406
Db	360	FGHQCLQGNLARAEMETALGTLFERFGLRLAAPADIPKPGSTIQGMLEPVTW	416

```

RECORD 9
US-09-861-289-39
; Sequence 39, Application US/09861289
; Patent No. US2520110897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/861.289
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-861-289-39

```

Query Match	16.7%;	Score 363.5;	DB 9;	Length 416;
Best Local Similarity	28.3%;	Pred. No. 1.le-27;		
Matches	97;	Conservative 61;	Mismatches 154;	Indels 21; Gaps 7;
76 QY	EDYRHSSECP	PIPERAGEAYDF	IPTSMDDPPEQ	RQFRALANQVGMPPVWDKLENRIQELA 135
76 DB	::: :::	::: :::	::: :::	::: :::
72 QY	KQWR--NSTT	PTAEAAALNHKX	LES--DPPRH	TRLRKLVAREETMERVELLRVQEI 127
72 DB	::: :::	::: :::	::: :::	::: :::
136 QY	CSLIESL--R	EOGQCNEFYD	AEPPPIRIFM	LLAGLPEEDIPLHKVLTDMCTPDCSMTF 193
136 DB	::: :::	::: :::	::: :::	::: :::
128 QY	DGLVDAMLA	APDGRALMES	LAWPIPTIV	ISELGVFEEDRAAFVWTDVAFPPDDPAQA 187
128 DB	::: :::	::: :::	::: :::	::: :::
194 QY	AEAKELYDY	IPIITEQRQK	EGTDAIS--IV	ANGQVNGRPITSCAEAKRCGLILVGLDT 252
194 DB	::: :::	::: :::	::: :::	::: :::
188 QY	QTAXAEMSG	YLSRLIDSKR	GQDGEDLLS	ALVETSDSDGSLTSEELGNHILVAGHET 247
188 DB	::: :::	::: :::	::: :::	::: :::
253 QY	VWNFLSFSE	WELAKSPER	HQELIEREPI	PACCELLRRFSIVAGRIITSDYEF--- 307
253 DB	::: :::	::: :::	::: :::	::: :::
248 QY	TVNIANGY	ALLSHDFC	LAALRADMT	LLDGAVEMLR-----YEGPVESATYRFPVEPW 302
248 DB	::: :::	::: :::	::: :::	::: :::
308 QY	--HGVLKKG	QDQI-LFQ	MSLSDEREN	ACPMHVDPSROKVSHTTTSHGSHLGLGQHARR 365
308 DB	::: :::	::: :::	::: :::	::: :::
303 QY	DLDGTVI	ZAGTVLV	LADAHRT	PERFPDPHFRDITRTAGHLAFSGHGFICIGAPLARL 362
303 DB	::: :::	::: :::	::: :::	::: :::
366 QY	EIIIVTLK	EWLIRI	TFDS--IAP	GAQIQHSGIVSGVQALPLVM 436
366 DB	::: :::	::: :::	::: :::	::: :::
363 QY	EARTAVRA	LLERCED	LADVSGEL	VWYVYENPMIRGLKALPIRW 405
363 DB	::: :::	::: :::	::: :::	::: :::

ESULT 10
S-09-860-846-39
Sequence 39, Application US/09860846

```

; Patent No. US20020164742A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
;
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
;
; FILE REFERENCE: 600.438US1
;
; CURRENT APPLICATION NUMBER: US/09/860,846
; CURRENT FILING DATE: 2001-05-18
;
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
;
; NUMBER OF SEQ ID NOS: 43
;
; SOFTWARE: FastSEQ for Windows Version 3.0
;
; SEQ ID NO 39
;
; LENGTH: 416
;
; TYPE: prt
;
; ORGANISM: Streptomyces venezuelae
;
; US-09-860-846-39

```

Query Match	16.4%	Score 363.5	DB 9	Length 416	
Best Local Similarity	28.3%	Prod. No. 1.1e-27			
Matches	97	Conservative 61	Mismatches 164	Indels 21	Gaps 7
QY	76	EDYRHSSBCPFFPRAGAYDIFPTSDPPEORQFALANQVGVVVKLENRIQELA	135		
DB	72	KQWR--NSTTETPEAAALNNHLES--DPPIRTIRKLVARFETMRVVELLRVQVEIV	127		
QY	136	CSLIESL--RPOQCNCFTEDYAEPPFIRIFMLLAGLPEEDIPIHLKYLTQDMTRPDGSMWF	193		
DB	128	DGLVDAMLAAPDGRADLMESLAWPLDITVISELLGWPEPDRAAFVWDFAVFPDDPAQA	187		
QY	194	AAKAKELNYLPIIIEORQKPGTDAIS--IVANGVNGRPITSCAKRMCGMLLVGSLDT	252		
DB	188	QTAAEMSGYLSLIDSKRGQGEDLLSALVTSDEGSGRLTSEELGMAHILLVAGHET	247		
QY	253	VYNLFSFSEELAKSPHEQELIERPERIPAACEEILLRRESLVADGRILTSOYEF----	307		
DB	248	YNLIANGMYALLSHPDQAAALRADMTLLDGAVEMLR-----YEGFEVSATYRFPVEPV	302		
QY	308	--HGVLKKGEDQILLQMLSGLDRENACPMEHVDFSRKVSHTTFGHGSHLCIGQHLAR	365		
DB	303	DLDGTVIPAGDTVLVVLADAHTPERFPDPRHFDIERTAGHLAFGHGIFHCIGAPLARL	362		
QY	366	ELIVTLKSWLRIPEFS--IAPCAQIOHKSIGVSGVQALPLVW	406		
DB	363	EARIARVALLERCPLQALDVSQGEVWYVNPMPMRIGLKALPIRW	405		

```

RESULT 11
US-09-988-384B-39
; Sequence 39, Application US/09988384B
; Publication No. US20030073824A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.536US1
; CURRENT APPLICATION NUMBER: US/09/988,384B
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: PCT/US99/14398
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-988-384B-39

```

Query Match	16.7%;	Score 363.5;	DS -0;	Length 416;
Best Local Similarity	28.3%;	Pred. No. 1.1e-27;		
Matches	97;	Conservative	61;	Mismatches 164; Indels 21; Gaps 7;
QY	76	BOYRHSCECFIPREAGEAIVCFPTSDPQRCFRALANQVGVMPVVKLENRIQELA	135	
Db	72	KDWR--NSTTPTAEAAALNNMUES--DPPHRTKRLKLVAREFTMARVELLRERVOEIV	127	
QY	136	CSLIESL--RFGQGCNFTEDYAEFPFPIRIFMLLAGLPEEDIPHLKYLTQDMTRDGGMTF	193	
Db	128	DGLVDMLAAEDGRADLMESLAWELPTVISELLGVPEDEAAFRVWTDARFFDDPAQA	187	
QY	194	AEAKEYLDYLIPIEORRQXPGGTALIS-IVANGQVGRPITSDEAKEMCGLLLVGSELT	252	
Db	188	QTAAEMSGYLSRLTDSKRGODGELLASLVTSDEDSRLTSEELGMAHILLVAGHET	247	
QY	253	VYNFLSFSEMEFLAKSPEHRQELIERPERIPAACELLARSLVADGRILLTSDYEF----	307	
Db	248	TVNLIANGMYALLSHPOOLAAALRADMTLLDGAVEEMLR-----YEGVESATRYFPVEPV	302	
QY	308	--HGVLQKKGQDILLPQMLGLDERENACPMHVDFSRKVSHSTTFHGSHLCGLQHLARR	365	
Db	303	DLDCGTVPAGDGTIVLVLAADAHRTFERFDEHRFDIROTAGHLAFGHIHFCIGAPLRL	362	
QY	366	EIIVTLKEWLTTRIDFS--TAPGAQIOHKSGIVSGVQALPLVW	406	
Db	363	EARIARVALLERCPLADYSPGLVWYVNPIMGALPLTRW	405	

```

RESULT 12
JS-09-836-821-39
; Sequence 39, Application US/09836821
; Publication No. US2003087405A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/836,821
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
JS-09-836-821-39

```

Query Match	-6.7%;	Score 363.5;	DB 10;	Length 416;	
Best Local Similarity	28.3%;	Pred. No. 11e-27;			
Matches	97;	Conservative 61;	Mismatches 164;	Indels 21;	Gaps 7;
QY	76	EYRHSSCPTIPRAGRAYDIFPTSMQPEQQRPRALANQVGVPPVVDKLENRIQELA	135		
Db	72	KDWR--NSTPLTEAAALNNHNLBS--DEPHRLRLKLVAREFTMRRVVELLRPRVQEIIV	127		
QY	136	CSLISL--RPOQCNTEDYAEPPPIRIFMLLAGLPDEEDIPHLKYLTDQMTRPDGSMTF	193		
Db	128	DELVDAMLAAPOGRADLMESLAWLPITVISELLGVPEPDRAAFRWTTDAFVFPDPAQA	187		
QY	194	AEAKALYLYLPIIPIEQRCKPGTDAIS--IVANGCVNREPIITSDRAKMGCLLLVGLGDT	252		
Db	188	QTAMEMSGYLSRLTOSKRGQGEDLLSALVTSDEDSRLTSELLGNMHTLLVAGHET	247		
QY	253	VYNFLSFMSEFLAKSPHQRQETIERPERIPAAACELLRRFSLVADGRILITSDYER----	307		
Db	248	TWNLIANGVYALLSHPDQLAALRADMT--LDGAVEMLR-----YGGPVESATIRPFPEIV	302		
QY	308	--GGVOLKKGDOILLPOMLSGLDERENACPMKWFDSRQKVSHTTTGGHSHLCAQHLARR	365		

```

Db      303 DLDCGVTPAGCTVVLWLAHAKTERPPDRFRCRDTAGHLAFGHGHIFCIGAPLRL 360Z
           :   : |||=: :       |     |    |   |   |: |   |   |||
Oy      366 EIIVTLKEWITRIDFS--IAPGAQIOHKSGIVSVCALPLVV 406
           |   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Dd      363 EARIAVRALLERCDFOLDVSPGELVWPENMIRGLKALPIRW 405
           |   ::|||::|||::|||::|||::|||::|||::|||::|||::|||

```

RESULT 13
US-09-793-708-18
; Sequence 18, Application US/09793708
; Publication No. JS20030104597A1
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: McDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002121
; CURRENT APPLICATION NUMBER: US/09/793,708
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 09/657,440
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 09/320,878
; PRIOR FILING DATE: 1999-05-27

```

; PRIOR APPLICATION NUMBER: US 09/073,538
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: US 08/946,247
; PRIOR FILING DATE: 1997-04-30
; PRIOR APPLICATION NUMBER: US 60/134,990
; PRIOR FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-793-708-18

Query Match          16.7%   Score 363.5;   DE 10;   Length 416;
Best Local Similarity 28.3%;   Pred. NO. 1.1e-27;
Matches 97;   Conservative 61;   Mismatches 164;   Indels 21;   Gaps 7;

QY      76  EDVRFHSESCRPPEAGAYDFIPSTMPEPPQRFALANQVGVHPVVDKLENRQELA 135
       111:  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      72  KQWR--NSTTPTTEAALNNHMLCS--DPPRHTLRKLIVAREFTMRVRVLLPRVQEI 127

```

```

128 DGLVDAIAAPGRADLMESLAWLPITVISLLGVPEPDRAAFVYWTDAFVPPDPAQA 187
194 ABAEAKALYDLIPIIQRQKPGTDAIS--IVANGQVNGRPITSDCAKRMCGLLLVGGLDT 252
198 QTAAEMSGYLSRLDSKRGQGBULLSALVTSDEGSLTSEELGMAHILLVAGHET 247
253 VYVNFUSFGEFLAKSPHEHQELIERPERIPAAACELLRRFSLVADGRILTSYDF----- 307
248 TVNLIANGVYALSHPDOLAAALRADMTLLDGAVEEMLR-----YEGPVSATYRFPVPV 302
308 --HGVLKKGDOILLPOMLSGLDERENACEPHVDFSRQKVSHTTFGHSHLCIGQHLARR 365
303 DLDGTVIPAGDTVLVVLDAHRTPRFFDPRHRFDTRDTAGHLAFGHGHICGAPLARL 362
366 ELIVTLKWLTRIPDFS--IAPGAQIOHKSGIVSGVQALPLVW 406
363 EARTAVRALERCPLADLVSPGELVWYYPNMIKGLKALPIRW 405

RESULT 4
US-1.5-201-365-13

```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 18:49:16 ; Search time 14.9588 Seconds
(without alignments)
1428.803 Million cell updates/sec

Title: US-09-246-451A-2
Perfect score: 2180
Sequence: 1 TTEIQQNANLAPPPHVPF.....IVSGVQALPFWDPATTYKAV 414

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pcp.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pcp.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pcp.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pcp.*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pcp.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	363.5	16.7	416	3	US-09-320-878-18
2	363.5	16.7	416	3	US-09-105-537-39
3	363.5	16.7	416	4	US-09-141-908-13
4	363.5	16.7	416	4	US-09-657-408-18
5	347	15.9	437	4	US-09-252-991A-17836
6	340	15.6	403	6	5212296-9
7	333	15.3	409	3	US-09-385-028-12
8	332	15.3	409	4	US-09-726-614-12
9	333	15.3	409	4	US-09-385-040-12
10	330	15.1	406	6	5212296-6
11	324.5	14.9	399	4	US-08-765-907A-10
12	316	14.5	412	1	US-09-102-863-11
13	316	14.5	412	5	PCT-US92-10885-11
14	271.5	12.5	419	3	US-09-335-409-8
15	271.5	12.5	419	3	US-09-413-814-71
16	271.5	12.5	419	4	US-09-568-102-8
17	271.5	12.5	419	4	US-09-567-969-8
18	271.5	12.5	419	4	US-09-568-480-8
19	271.5	12.5	419	4	US-09-568-486-8
20	271.5	12.5	419	4	US-09-568-472-8
21	271.5	12.5	419	4	US-07-899-8
22	261	12.0	395	4	US-09-266-965-129
23	233	10.7	468	4	US-09-252-991A-32437
24	169.5	7.8	189	4	US-09-679-279-20
25	161	7.4	422	2	US-09-096-982-5
26	161	7.4	422	2	US-08-653-650A-5
27	161	7.4	474	2	US-09-096-982-8

28	161	7.4	474	2	US-08-653-650A-8	Sequence 8, Appli
29	158	7.2	443	2	US-09-096-982-9	Sequence 9, Appli
30	158	7.2	443	2	US-08-653-650A-9	Sequence 9, Appli
31	156	7.2	422	1	US-08-396-218-2	Sequence 2, Appli
32	156	7.2	422	1	US-08-760-116-2	Sequence 2, Appli
33	129	5.9	512	2	US-08-194-981B-5	Sequence 5, Appli
34	127.5	5.8	503	4	US-09-583-447A-2	Sequence 2, Appli
35	127	5.8	504	4	US-09-583-447A-4	Sequence 2, Appli
36	127	5.8	516	4	US-09-215-694-16	Sequence 16, Appli
37	126	5.8	503	4	US-09-144-367-2	Sequence 2, Appli
38	123.5	5.7	382	3	US-09-320-878-7	Sequence 7, Appli
39	123.5	5.7	382	3	US-09-141-908-7	Sequence 7, Appli
40	123.5	5.7	382	4	US-09-657-440-7	Sequence 7, Appli
41	123.5	5.7	402	3	US-09-105-537-22	Sequence 22, Appli
42	123.5	5.7	3782	3	US-09-105-537-4	Sequence 4, Appli
43	120.5	5.5	524	4	US-09-126-420A-24	Sequence 24, Appli
44	118	5.4	513	3	US-08-948-564-6	Sequence 6, Appli
45	115.5	5.3	490	1	US-08-201-118-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-320-878--8
; Sequence 18, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: McDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT CARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320.878A
; EARLIER FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141.908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073.538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846.247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119,139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100,880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-320-878--8

Query Match 16.7% Score 363.5; DB 3; Length 416;
Best Local Similarity 28.3%; Pred. No. 3.4e-29;
Matches 9; Conservative 61; Mismatches 164; Indels 21; Gaps 7;
QY 76 EDYHSESCEFFIPREAGEAYDFIPTSMDFEQQRALANQVGVMPVVKLENIQSLA 135
Db 72 KDWK--NSTPLTEAEALNHNMLES--DFPHTRUKLVAREFTMRVLLRPRVQEV 127
QY 136 CSLIESI--RPQCQNFTEYAPFPPIRIFMLAGIPEEDIPHLKYLTDOMTRPDGSM 193
Db 128 DGLVDLAMLAPDGRADIMESLAWPLPITVISSELLGVPEPDRAAFRVWTFDVFDDPAQA 187
QY 194 AEAKEALYDLIPI--RORROKPCPTDAIS--IVANQVNGRPITSCAEKRMCGLLIVGLDT 252
Db 188 QTAMMSGYLSRLIDSKRGQDGEDLLSALVRTSDEGSRLTSEELGMAHILLVAGHET 247

```

QY 253 VVNFSLFSEMFIAKSPHROELIERPERIPAAACEELIRFSLVADGRILTSDYEF----- 307
Db 248 TVNLIANGMYALLSHPOQALRADMTLLDGAVEMLR-----YEGPVESATYRFPVEPV 302
QY 308 --HGVLKKGQDQILLPQMLSGLDRENAACPMHVDPSRQKVSHHTFGHSHLCLGQHLARR 365
Db 303 DLDTGTVIPAGDTVLVVLADAHRTPERFPDHRFDIRDTAGHLAFGHGHCIGAPLARR 362
QY 366 EIIIVTKEMWLRIPDFS--IAPGAQIQHKSGLVSGVQALPLVW 406
Db 363 EARIAVRALLERCPDLALDVSPGELVWYFNPMPMRGLKALPIRW 405

```

RESULT 2

```

US-09-105-537-39
; Sequence 39, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, P. H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-105-537-39

```

```

Query Match 16.7%; Score 363.5; DB 3; Length 416;
Best Local Similarity 28.3%; Pred. No. 3.4e-29;
Matches 97; Conservative 61; Mismatches 164; Indels 21; Gaps 7;

```

```

QY 76 EDYRHSSECPPIPREAGEAYDPTSMDDPEQRFALANQVGMVVDKLENRIQELA 135
Db 72 KQWR--NSTPLTEAEALNHNMLE--DPPRHLRLKLVAREFTMRVRLPRVQEIIV 127
QY 136 CSLIESL--RPQGCNFTEDYAEPPRIRIFMLAGLSEEDIPHLKYLTDQTRPDGSMTF 193
Db 128 DGLVDAMLAPDGRADLMESLAWPLPTVITSELLGVPEPDRAAFRVWTDVFPDDPAQA 187
QY 194 AEAKEALYDLIPIEOROKPGTDALS-IVANGQVNGRPIITSDEAKRMCGILLVGLDT 252
Db 188 QTAMAEWSGYLSRLDSKRGQDGLLSALVRTSDEGSLTSELLGMAHILLVAGHET 247
QY 253 VVNFSLFSEMFIAKSPHROELIERPERIPAAACEELIRFSLVADGRILTSDYEF----- 307
Db 248 TVNLIANGMYALLSHPOQALRADMTLLDGAVEMLR-----YEGPVESATYRFPVEPV 302
QY 308 --HGVLKKGQDQILLPQMLSGLDRENAACPMHVDPSRQKVSHHTFGHSHLCLGQHLARR 365
Db 303 DLDTGTVIPAGDTVLVVLADAHRTPERFPDHRFDIRDTAGHLAFGHGHCIGAPLARR 362
QY 366 EIIIVTKEMWLRIPDFS--IAPGAQIQHKSGLVSGVQALPLVW 406
Db 363 EARIAVRALLERCPDLALDVSPGELVWYFNPMPMRGLKALPIRW 405

```

RESULT 3

```

US-09-141-908-13
; Sequence 13, Application US/09141508
; Patent No. 6503741
; GENERAL INFORMATION:
; APPLICANT: Ashley, Gary
; APPLICANT: Betlach, Melanie C.
; APPLICANT: Betlach, Mary
; APPLICANT: McDaniel, Robert
; APPLICANT: Tang, Li

```

```

; TITLE OF INVENTION: Combinatorial Polyketide Libraries Produced Using a
; TITLE OF INVENTION: Modular PKS Gene Cluster as Scaffold
; FILE REFERENCE: 300622002100
; CURRENT APPLICATION NUMBER: US/09/141,908
; CURRENT FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/373,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: PROV. 60/076,919
; EARLIER FILING DATE: 1998-03-05
; EARLIER APPLICATION NUMBER: PROV. 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-141-908-13

```

Query Match 16.7%; Score 363.5; DB 4; Length 416;

```

Best Local Similarity 28.3%; Pred. No. 3.4e-29;
Matches 97; Conservative 61; Mismatches 164; Indels 21; Gaps 7;
QY 76 EDYRHSSECPPIPREAGEAYDPTSMDDPEQRFALANQVGMVVDKLENRIQELA 135
Db 72 KQWR--NSTPLTEAEALNHNMLE--DPPRHLRLKLVAREFTMRVRLPRVQEIIV 127
QY 136 CSLIESL--RPQGCNFTEDYAEPPRIRIFMLAGLSEEDIPHLKYLTDQTRPDGSMTF 193
Db 128 DGLVDAMLAPDGRADLMESLAWPLPTVITSELLGVPEPDRAAFRVWTDVFPDDPAQA 187
QY 194 AEAKEALYDLIPIEOROKPGTDALS-IVANGQVNGRPIITSDEAKRMCGILLVGLDT 252
Db 188 QTAMAEWSGYLSRLDSKRGQDGLLSALVRTSDEGSLTSELLGMAHILLVAGHET 247
QY 253 VVNFSLFSEMFIAKSPHROELIERPERIPAAACEELIRFSLVADGRILTSDYEF----- 307
Db 248 TVNLIANGMYALLSHPOQALRADMTLLDGAVEMLR-----YEGPVESATYRFPVEPV 302
QY 308 --HGVLKKGQDQILLPQMLSGLDRENAACPMHVDPSRQKVSHHTFGHSHLCLGQHLARR 365
Db 303 DLDTGTVIPAGDTVLVVLADAHRTPERFPDHRFDIRDTAGHLAFGHGHCIGAPLARR 362
QY 366 EIIIVTKEMWLRIPDFS--IAPGAQIQHKSGLVSGVQALPLVW 406
Db 363 EARIAVRALLERCPDLALDVSPGELVWYFNPMPMRGLKALPIRW 405

```

RESULT 4

```

US-09-657-440-18
; Sequence 18, Application US/09657440
; Patent No. 6509455
; GENERAL INFORMATION:
; APPLICANT: Ashley, Gary
; APPLICANT: Betlach, Melanie C.
; APPLICANT: Betlach, Mary C.
; APPLICANT: McDaniel, Robert
; APPLICANT: Tang, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/657,440
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/320,878
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: CIP OF 09/141,908
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 416
; TYPE: PRT

```


db 229 GLVQAADESGQLSEAEVLSMAHLLMMSGFETTMIGNALVTLLVNPEQLALLRAQPELL 288

db 229 GLVQAADESGQLSEAEELVSMHLLMMSGFEETMNMIGNALVTLLVNPEQLALRAQPELL 288

RESULT 7
US-09-385-028-12
; Sequence 12, Application US/09385028
; Patent No. 6232106
; GENERAL INFORMATION:

APPLICANT: Susan E. Jensen
 APPLICANT: Kwamena A. Aidoo
 APPLICANT: Ashish S. Paradkar
 TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic
 Patent No. 6232106
 TITLE OF INVENTION: Acid Biosynthesis
 NUMBER OF SEQUENCES: 25
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN PPLC
 STREET: The Jenifer Building, 400 Seventh Street, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/395,028
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/790,462
 FILING DATE: 29-JAN-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: D. Douglas Price
 REGISTRATION NUMBER: 24,514
 REFERENCE/DOCKET NUMBER: 1418/P57452US2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 638-6666
 TELEFAX: (202) 39305350
 TELEX: RCA 248593 IDEA UR
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 409 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-385-028-12

Query Match 15.3%; Score 333; DB 3; Length 409;
 Best Local Similarity 29.3%; Pred. No. 5e-26;
 Matches 122; Conservative 60; Mismatches 185; Indels 54; Gaps 18;

QY 16 PHVPEHLVDFDMYNSLSAGVQEAVALQESNVPDLVWTRCNGGH-WIATRGQLIREA 74
 DB 14 PAYPMRVCPVD---PPQLAGLSRQKAASRVT---LW---DGSQVWLVTSAGARAV 62
 QY 75 YEDYRHFS-SECFPIP-----REAGEAYDFIPTSDPPRQRFRA-----LANQV 118
 DB 63 LGDRRRTAVTSAPGFPMLTRTSOLVRANPESASFI--RMDDPQHSRLSMLTRDFLAR 120
 QY 119 VGM-PVWDXLKNRIQELACLSTIESLPQGCNFTEDYAEFPPIRIFMLAGLPEEDIPHL 177
 DB 121 EALRPFAVREL---LDEILGGLVKGERP---VLDVAGLTTPVPSRVITLLFGAGDDRREFI 174
 QY 178 K----YLTQMTPTDGSMTFAEAKALDYLLPIEQRKQKPGTDAISIVANGQVNGRPI 233
 DB 175 EDRSAVLIDRGTYPE---QVAKARDLDGLRELVEIERENFG--ELISRLVIDQVRPGHL 231
 QY 234 TSDEAKRMGLLIVGLDVTWVNFSLSPSMEFLAKSPERQELIERPERIPAAACELLRRFS 293
 DB 232 RVEENVEMCRLLIVAGHGTTTQASLSLLTDFLAGRLIEDPALLPAVEELLRRFS 291
 QY 294 LVADG--RLTSCYSEHGVLKGGDQILLPQMLSGLDERENACEMHVSFRQKVSHTTFG 351
 DB 292 IVQNGLARAAVDVQLDDLIBAGEGWLISLGNRDETVPFDPDVRDVRDARRHAFG 351
 QY 352 HGSHCLGGLHAR---RELIVLKEWLTRIPFSIA-PGAQOQHKSGIVS-GVQALPLVW 406

Db 352 HGMCQCLQWLARVELEILAAVLRMM---PGARLAVPPELDFRHEVSSYGLGALPVTW 408
 QY 407 D 407
 Db 409 Z 409
 RESULT 8
 US-09-726-614-12
 Sequence 12, Application US/09726614
 Patent No. 6514735
 GENERAL INFORMATION:
 APPLICANT: Susan E. Jensen
 APPLICANT: Kwamena A. Aidoo
 APPLICANT: Ashish S. Paradkar
 TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic
 Patent No. 6514735
 TITLE OF INVENTION: Acid Biosynthesis
 NUMBER OF SEQUENCES: 25
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN PPLC
 STREET: The Jenifer Building, 400 Seventh Street, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09726,614
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/790,462
 FILING DATE: 29-JAN-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: D. Douglas Price
 REGISTRATION NUMBER: 24,514
 REFERENCE/DOCKET NUMBER: 1418/P57452US2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 638-6666
 TELEFAX: (202) 39305350
 TELEX: RCA 248593 IDEA UR
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 409 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-726-614-12

Query Match 15.3%; Score 333; DB 4; Length 409;
 Best Local Similarity 29.3%; Pred. No. 5e-26;
 Matches 122; Conservative 60; Mismatches 185; Indels 54; Gaps 18;

QY 16 PHVPEHLVDFDMYNSLSAGVQEAVALQESNVPDLVWTRCNGGH-WIATRGQLIREA 74
 DB 14 PAYPMRVCPVD---PPQLAGLSRQKAASRVT---LW---DGSQVWLVTSAGARAV 62
 QY 75 YEDYRHFS-SECFPIP-----REAGEAYDFIPTSDPPRQRFRA-----LANQV 118
 DB 63 LGDRRRTAVTSAPGFPMLTRTSOLVRANPESASFI--RMDDPQHSRLSMLTRDFLAR 120
 QY 119 VGM-PVWDXLKNRIQELACLSTIESLPQGCNFTEDYAEFPPIRIFMLAGLPEEDIPHL 177
 DB 121 EALRPFAVREL---LDEILGGLVKGERP---VLDVAGLTTPVPSRVITLLFGAGDDRREFI 174
 QY 178 K----YLTQMTPTDGSMTFAEAKALDYLLPIEQRKQKPGTDAISIVANGQVNGRPI 233

Db 175 EDRSAVLDRGYTPE---QVAKARDELGYLRELVDRIENPGTDLISRLVIDQVRPGHL 231
QY 234 TSDEAKRMCGLLLVGLDVTNNFLSFSMEFLAKSPEHRQELTERPERIPAKCEELLRRFS 293
Db 232 RVEEMVPCRELLLVAGHGTTTSQASLSLSLTDELAGRUTEDPALLPKAVEELLRRHS 291
QY 294 LVADG--RLTSDYEFHGVLKKGQOILLPQMLSGLDRENAACPMHVPFSROKVSHTTFFG 351
Db 292 IVQNGLAARAAVEDVQ-LDVLRAGEGVLSLSAGNRDETVPDPORVDVDRARRHLAFG 351
QY 352 HGSHLCLGQHLAR---REIIVT-KWLTTRIPDFSIA-PCAQIQHKSGLVSV-GVQALPLVW 406
Db 352 HGMEQCLGQHLARVELEELAAVLKWM---PGARLAVPFEELDFRHEVSSYGLGALPVTW 408
QY 407 D 407
Db 409 Z 409

RESULT 9
US-09-385-040-12
; Sequence 12, Application US/09385040
; Patent No. 6589775
; GENERAL INFORMATION:
; APPLICANT: Jensen, Susan E
; APPLICANT: Aicoo, Kwamena A
; APPLICANT: Paradkar, Ashish S
; TITLE OF INVENTION: DNA SEQUENCE ENCODING ENZYMES OF CLAVALANTIC ACID
; FILE REFERENCE: 09/385,040
; CURRENT APPLICATION NUMBER: US/09/385,040
; PRIOR FILING DATE: 1999-08-30
; PRIOR APPLICATION NUMBER: US 08/790,462
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 12
; TYPE: PRT
; ORGANISM: Streptomyces clavuligerus
US-09-385-040-12

Query Match 15.3%; Score 333; DB 4; Length 409;
Best Local Similarity 29.0%; Pred. No. 5e-26;
Matches 122; Conservative 60; Mismatches 185; Indels 54; Gaps 18;

QY 16 PHVPEHLWFDQMYNPSNLSAGVQEAVALICSNVDELVWTRCNGGH-WIATRGQLIREA 74
Db 14 PAYPHERVCPVD---PPQLAGLRSSQKAASRVT-----LW---DGSQVMLVTSHAGARAV 62
QY 75 YEDYRHS-SECFPIF-----REAGEAYDPIPTSDPPPEQRCFRA-----LANOV 118
Db 63 LGDRRTAVTSAPGPEMLRTSCLVEANPESASFI--RNDDPQHSRLRSMLTRDFLARRA 120
QY 119 VGM-PWVOKLENIQELACLSIESLRPQCCNFTEDYAEFPPIRIFMLLAGLFEEDIPHL 177
Db 121 EALRAVAVREL---LDEILGLLVKGRP---VDLVAGTTPVPSRVITLLFGADDDRREFI 174
QY 178 K----YLTDMQTRPDGSMTFEAKKALYDYLPIIPIORQKPGTDAISIVANGVNGRPI 233
Db 175 EDRSAVLDRGYTPE---QVAKARDELGYLRELVDRIENPGTDLISRLVIDQVRPGHL 231
QY 234 TSDEAKRMCGLLLVGLDVTNNFLSFSMEFLAKSPEHRQELTERPERIPAKCEELLRRFS 293
Db 232 RVEEMVPCRELLLVAGHGTTTSQASLSLSLTDELAGRUTEDPALLPKAVEELLRRHS 291
QY 294 LVADG--RLTSDYEFHGVLKKGQOILLPQMLSGLDRENAACPMHVPFSROKVSHTTFFG 351
Db 292 IVQNGLAARAAVEDVQ-LDVLRAGEGVLSLSAGNRDETVPDPORVDVDRARRHLAFG 351
QY 352 HGSHLCLGQHLAR---REIIVT-KWLTTRIPDFSIA-PCAQIQHKSGLVSV-GVQALPLVW 406
Db 352 HGMEQCLGQHLARVELEELAAVLKWM---PGARLAVPFEELDFRHEVSSYGLGALPVTW 408

QY 407 D 407
Db 409 Z 409

RESULT 10
5212296-6
; Patent No. 5212296
; APPLICANT: DEAN, CAROLINE; HARDER, PATRICIA A.; LETO, KENNETH
; J.; O'KEEFE, DANIEL P.; OMER, CHARLES A.; ROMESSER, JAMES A.
; TEPPEMAN, JAMES M.
; TITLE OF INVENTION: EXPRESSION OF HERBICIDE METABOLIZING
; CYTOCHROMES
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/569,781
; FILING DATE: 23-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 464,499
; FILING DATE: 12-JAN-1990
; APPLICATION NUMBER: 405,605
; FILING DATE: 11-SEP-1989
; SEQ ID NO: 6;
; LENGTH: 406
5212296-6

Query Match 15.1%; Score 330; DB 6; Length 406;
Best Local Similarity 26.3%; Pred. No. 1e-25;
Matches 104; Conservative 74; Mismatches 193; Indels 24; Gaps 11;

QY 31 PSNLSAGVQ--EAWAVIQESNVPLVTRCNGGH-WIATRGQLIREAYEDYR----- 79
Db 17 PSNRSCPQLPDGVAQLRDTGPHLRVTLYDGRQAWVTKEAARKLIGDPLSSNRD 76
QY 80 HFSECCPIF--REAGEAYDPIPTSDPPPEQORFALANQVGMVVDKLENRIQELACS 137
Db 77 RFATSPRFRAVRESPOAF-----IGLDPPEHGTTRRTTISEFTVARIKGMPEVEVHG 132
QY 138 LIESLRPQG-QCNFTEDYAEFPPIRIFMLLAGLFEEDIPHLKYLTDMQTRPDGSMTFEAE 196
Db 133 FLDEMLAAGTADIVSQFALPVPVSWVICLLGVYADHEFFQDASKRLVQSTDAQSALTA 192
QY 197 KEALYDYLPIIPIORQKPGTDAI-SIVANGVNGRPIITSDAEKRMCGLLLVAGGLDVTVN 255
Db 193 RNDLAGYLDGLITQFQTEPGAGLVGALVADQLANGE-IDREELISTAMLLIAGHETTAS 251
QY 256 FLFSMEFLAKSPEHRQELTERPERIPAKCEELLRRFSL--VADGRILTSDFYEHGVOLK 313
Db 252 MTSLSVITLLDHPQYAAALRADSLVPGAVEELLRYLAADIAGGRVATADIEVEGHLIR 311
QY 314 KGOQILLPQMLSGLDRENAACPMHVPFSROKVSHTTFFHGSHLCLGQHLARREIIVTLKE 373
Db 312 AGEGVIVANSIARNDKGVYEDPDALDIHRSARHHLAFGFGVHQCGLGONLARLEVLINA 371
QY 374 WLTRIPDFSIA-PCAQIQHKSGLVSV-GVQALPLVW 406
Db 372 LMDRVETLRAVPEVQLVLRPGTTIQGVNELPVTW 406

RESULT 11
US-08-765-907A-10
; Sequence 10, Application US/08765907A
; Patent No. 6352839
; GENERAL INFORMATION:
; APPLICANT: BLANC, Veronique
; APPLICANT: THIBAUT, Denis
; APPLICANT: BAKAS-JACQUES, Nathalie
; APPLICANT: BLANCHE, Francis
; APPLICANT: COUZET, Joel
; APPLICANT: BARRIERE, Jean-Claude
; APPLICANT: DEBUSSCHE, Laurent
; APPLICANT: FAMECHON, Alain

COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh System, 6.0
SOFTWARE: Microsoft Word, 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/10885
FILING DATE: 19921226
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: GALLEGOS, R. THOMAS
REGISTRATION NUMBER: 32,692
REFERENCE/DOCKET NUMBER: CR-9000-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-7342
TELEFAX: 302-892-7949
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 412 amino acids
TYPE: AMINO ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
PCT-US92-10885-11

Query Match 14.5%; Score 315; DB 5; Length 412;
Best Local Similarity 27.2%; Pred. No. 3e-24;
Matches 84; Conservative 56; Mismatches 163; Indels 6; Gaps 5;

QY 103 MDPPEQRFALANQVGMVVDKLENRIQELACSLIESLRQCG-QCNFTEDYAEPPPIR 161
Db 105 VDFEHNTRQRMILPTFSVKRIGALRPRIQTVDRLLDAMERQGPFAELVSAFALVPFSM 164
QY 162 IFMLLAGLFEEDIHLKYLTPQWTRPDGSMTFABAKEALDYLIPIIQRCKKPGCTDAIS 221
Db 165 VICALLGVYADHAFFERSQRLLEGPCADQVNRARDEBELGALIDRKRAEPDGLLD 224
QY 222 IVANGVNGRPIITDEAKMCGLLLVGLDITVWFLSFMSFLAKSPHROELIERPERI 281
Db 225 ELIRDRHEDGVDRQCLVAFVAVILLIAGHETTANNISLGTFTLLSHPEQLAALRAGGTST 284
QY 282 PAACEELLRRFSVADG--RLTSDYEFHGVQLKGDQDILLPQMLSGLDERENACPMHVD 339
Db 285 AVVTEELL-RLFSAEGLQRLATDMEVDGATIRKGGVWFSTSLINRDADVFPRAETLD 343
QY 340 FSRQKVSHTTGHGSHLCLGQLARRELIIVTLKEWLTTRIPDFSIA-PCAQ-CHKSG-IVS 397
Db 344 WDRPARHLAFGFGVHQLGQNLARAEIDIAMRTLFEPLGSLAVPAHEIRHKGDTIQ 403
QY 398 GVQALPLVW 406
Db 404 GLLDLPVAV 412

RESULT 14
US-09-335-409-8
; Sequence 8, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335,409
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Sorangium cellulosum

US-09-335-409-8

Query Match 12.5%; Score 271.5; DB 3; Length 419;
Best Local Similarity 23.8%; Pred. No. 1.4e-19;
Matches 100; Conservative 67; Mismatches 179; Indels 75; Gaps 14;

QY 6 QSNANLAPPHVPEHLVFFDFMYPNSLSAGVQE-----AWAVLQESNVDPDLVWTRCNGG 61
Db 3 QEQANQSETKP-----AFDFKFPAP-----GVAEDPFPALERLREA-TPIFYWD--EGR 48
QY 62 HMLATSEQ-----LIREAYEDYRHFSSBCEPFPREAGEAYDFTPTSMDDPEQR 109
Db 49 SWLTRYHDVSAVERDERFAVSEEWESAEYSSAIZ-----ELSDMKYGLFGLPPSDHA 104
QY 110 QFRALANQVGMVVDKLENRIQELACSLIESLRQCGQCNFTEDYAEPPPIRIFMLLAGL 169
Db 105 RYRKLVNPSFTSRAIDLLRAEIQRTVDQLDARSQGEFDFVVDYAEGLPMRAISALLKV 164
QY 170 PREDIPHLKYLTDQWTRPDGSMTFABAKEALDYLIPI-----I 207
Db 165 PAE-----CDEKFRREGSAT-----ARALGVGLVPEQVDEETKTLVASVTEGLALLHDV 212
QY 208 IFORQKQ2-GTDAISIVANGVNGRPIITSDAEKRCNCELLVGLDITVWFLSFMFLAK 266
Db 213 LDERRNPLENDVLTMLLQAEADGSRSLSTKELVALGVAIIAAGTDTTIIYIAFAVLNLLR 272
QY 267 SPEHRQELIERPERIPAAACEELLRRFSVADG--RLTSDYEFHGVQLKGDQI---LJFQ 322
Db 273 SPEALELVKAEPLMKNALDEVLRFDNI-RLITVRFARQDLEYCGASIKKGEWVFLIPS 332
QY 323 MSLGDERENACPMHVDPSRQKVSHTTGHGSHLCLGQLARRELIIVTLKEWLTTRIPDFS 382
Db 333 ALR--DGTVFSREDVDFVDRDGTASLAYGRGPHVCGVSLARLEAEIAGTIFRREPENK 390
QY 383 I 383
Db 391 L 391

RESULT 15
US-09-413-814-71
; Sequence 71, Application US/09413814
; Patent No. 6225664
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bloeker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hefle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
; TITLE OF INVENTION: heteropolyketide compounds
; FILE REFERENCE: PCT/US 99/23535
; CURRENT APPLICATION NUMBER: US/09/413,814
; CURRENT FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; EARLIER FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 71
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-413-814-71

Query Match 12.5%; Score 271.5; DB 3; Length 419;
Best Local Similarity 23.8%; Pred. No. 1.4e-19;
Matches 100; Conservative 67; Mismatches 179; Indels 75; Gaps 14;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 18:53:57 ; Search time 14.974 Seconds

(without alignments)

3108.883 Million cell updates/sec

Title: US-09-246-451A-11

Perfect score: 2180

Sequence: 1 ITTETQSNANLAPPHVPE.....IVSGVQALPLYWDPAITTKAV 414

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 107772 seqs, 262633353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2180	100.0	414	15	US-10-453-104-11
2	2176	99.8	414	15	US-10-453-104-2
3	2173	99.7	414	15	US-10-453-104-12
4	2169	99.5	414	15	US-10-453-104-13
5	377.5	17.3	404	14	US-10-214-446-50
6	376.5	17.3	404	14	US-10-214-446-40
7	372	17.1	399	14	US-10-314-657-37
8	367	16.8	416	14	US-10-156-761-14997
9	365.5	16.8	416	9	US-09-861-289-39
10	365.5	16.8	416	9	US-09-860-846-39
11	365.5	16.8	416	10	US-09-988-3843-39
12	365.5	16.8	416	10	US-09-836-821-39
13	365.5	16.8	416	10	US-09-793-708-18
14	365.5	16.8	416	14	US-10-201-365-13
15	365.5	16.8	416	14	US-10-160-539-18

16	365.5	16.8	416	14	US-10-271-889-39	Sequence 39, Appl
17	363.5	16.7	411	14	US-10-156-761-8376	Sequence 8376, Ap
18	358	16.4	393	14	US-10-156-761-9525	Sequence 9525, Ap
19	357	16.4	405	14	US-10-214-446-38	Sequence 20, Appl
20	356	16.3	425	14	US-10-214-446-20	Sequence 20, Appl
21	354.5	16.3	457	14	US-10-156-761-11073	Sequence 11073, A
22	351.5	16.1	399	14	US-10-156-761-9914	Sequence 9914, Ap
23	350	16.1	418	12	US-10-389-647-559	Sequence 559, App
24	341.5	15.7	388	14	US-10-156-761-13776	Sequence 13776, A
25	338.5	15.5	408	14	US-10-214-446-4	Sequence 4, Appli
26	333	15.3	409	15	US-10-458-201-12	Sequence 12, Appl
27	332	15.2	392	14	US-10-214-446-32	Sequence 32, Appl
28	328.5	15.1	406	14	US-10-214-446-2	Sequence 2, Appli
29	322	14.8	401	14	US-10-156-761-8710	Sequence 8710, Ap
30	321	14.7	404	14	US-10-214-446-16	Sequence 16, Appl
31	318.5	14.6	396	14	US-10-214-446-14	Sequence 14, Appl
32	318	14.6	428	14	US-10-201-213-6	Sequence 6, Appli
33	317.5	14.6	404	14	US-10-156-761-14659	Sequence 14659, A
34	315	14.4	412	14	US-10-214-446-36	Sequence 36, Appl
35	315	14.4	421	14	US-10-156-761-9703	Sequence 9703, Ap
36	314	14.4	418	14	US-10-214-446-22	Sequence 22, Appl
37	307	14.1	421	14	US-10-214-446-18	Sequence 18, Appl
38	304.5	14.0	404	14	US-10-156-761-10431	Sequence 10431, A
39	302.5	13.9	415	14	US-10-214-446-46	Sequence 46, Appl
40	301.5	13.8	400	14	US-10-314-657-62	Sequence 62, Appl
41	298.5	13.7	402	14	US-10-205-032-8	Sequence 8, Appli
42	298	13.7	475	14	US-10-145-415-22	Sequence 22, Appl
43	297	13.6	470	14	US-10-145-415-6	Sequence 6, Appli
44	295.5	13.6	430	9	US-09-738-626-4117	Sequence 4117, Ap
45	295	13.5	399	14	US-10-156-761-7959	Sequence 7959, Ap

ALIGNMENTS

RESULT 1
US-10-453-104-11
; Sequence 11, Application US/10453:04
; Publication No. US20030207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Hyun Joo
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
; FILE REFERENCE: 4058/1827-US3
; CURRENT APPLICATION NUMBER: US/10/453.104
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 09/661,093
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 09/246,451
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/096,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/106,834
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant M7-4H
US-10-453-104-11

Query Match 100.0%; Score 2180; DB 15; Length 414;
Best Local Similarity 100.0%; Pred. No. 7,1e-212;
Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TTETIQSNANLAPLPHVPEHLVDFDMYNSNLGAGQEAVALQESNVDPDLVWTRCNG 60
DB 1 TTETIQSNANLAPLPHVPEHLVDFDMYNSNLGAGQEAVALQESNVDPDLVWTRCNG 60
QY 61 GHWIATRGQILREAYEDYRHFSSCEPPIPREAGEAYDFIPTSMDFPQORFRALANQVVG 120
DB 61 GHWIATRGQILREAYEDYRHFSSCEPPIPREAGEAYDFIPTSMDFPQORFRALANQVVG 120
QY 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYL 180
DB 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYL 180
QY 181 TDQWTRPDGSMTEFAEAKAALYDYLPIIEQRORKEGTDAISIVANGQVNGRPITSDAKR 240
DB 181 TDQWTRPDGSMTEFAEAKAALYDYLPIIEQRORKEGTDAISIVANGQVNGRPITSDAKR 240
QY 241 MCGLLLVGGGLDVTNNFLSFSMEFLAKSPEHRQELIERPERIPAAACELLRRFSLVADGRI 300
DB 241 MCGLLLVGGGLDVTNNFLSFSMEFLAKSPEHRQELIERPERIPAAACELLRRFSLVADGRI 300
QY 301 LTSDFEFGVQLKGGDQILLPOMLSGLDERKNACPMHVDFSRQKVSHTTFGSGSHLCLGQ 360
DB 301 LTSDFEFGVQLKGGDQILLPOMLSGLDERKNACPMHVDFSRQKVSHTTFGSGSHLCLGQ 360
QY 361 HLARREIIVTLKEWLTRIPDFSAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414
DB 361 HLARREIIVTLKEWLTRIPDFSAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414

```

RESULT 2

```

US-10-453-104-2
; Sequence 2, Application US/10453104
; Publication No. US20030207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Hyun Joo
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
; FILE REFERENCE: 4058/1E827-US3
; CURRENT APPLICATION NUMBER: US 10/453,104
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 09/661,093
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 09/246,451
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/086,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/106,834
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 414
; TYPE: PRT
; ORGANISM: P. Putida
; OTHER INFORMATION: Mutant M7-6H
US-10-453-104-2

```

```

Query Match 99.8%; Score 2176; DB 15; Length 414;
Best Local Similarity 99.8%; Pred. No. 1.8e-211;
Matches 413; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TTETIQSNANLAPLPHVPEHLVDFDMYNSNLGAGQEAVALQESNVDPDLVWTRCNG 60
DB 1 TTETIQSNANLAPLPHVPEHLVDFDMYNSNLGAGQEAVALQESNVDPDLVWTRCNG 60
QY 61 GHWIATRGQILREAYEDYRHFSSCEPPIPREAGEAYDFIPTSMDFPQORFRALANQVVG 120
DB 61 GHWIATRGQILREAYEDYRHFSSCEPPIPREAGEAYDFIPTSMDFPQORFRALANQVVG 120

```

```

QY 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYL 180
DB 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYL 180
QY 181 TDQWTRPDGSMTEFAEAKAALYDYLPIIEQRORKEGTDAISIVANGQVNGRPITSDAKR 240
DB 181 TDQWTRPDGSMTEFAEAKAALYDYLPIIEQRORKEGTDAISIVANGQVNGRPITSDAKR 240
QY 241 MCGLLLVGGGLDVTNNFLSFSMEFLAKSPEHRQELIERPERIPAAACELLRRFSLVADGRI 300
DB 241 MCGLLLVGGGLDVTNNFLSFSMEFLAKSPEHRQELIERPERIPAAACELLRRFSLVADGRI 300
QY 301 LTSDFEFGVQLKGGDQILLPOMLSGLDERKNACPMHVDFSRQKVSHTTFGSGSHLCLGQ 360
DB 301 LTSDFEFGVQLKGGDQILLPOMLSGLDERKNACPMHVDFSRQKVSHTTFGSGSHLCLGQ 360
QY 361 HLARREIIVTLKEWLTRIPDFSAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414
DB 361 HLARREIIVTLKEWLTRIPDFSAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414

```

RESULT 3

```

US-10-453-104-12
; Sequence 12, Application US/10453104
; Publication No. US20030207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Hyun Joo
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
; FILE REFERENCE: 4058/1E827-US3
; CURRENT APPLICATION NUMBER: US 10/453,104
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 09/661,093
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 09/246,451
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/086,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/106,834
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant M7-6H
US-10-453-104-12

```

```

Query Match 99.7%; Score 2173; DB 15; Length 414;
Best Local Similarity 99.8%; Pred. No. 3.6e-211;
Matches 413; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 TTETIQSNANLAPLPHVPEHLVDFDMYNSNLGAGQEAVALQESNVDPDLVWTRCNG 60
DB 1 TTETIQSNANLAPLPHVPEHLVDFDMYNSNLGAGQEAVALQESNVDPDLVWTRCNG 60
QY 61 GHWIATRGQILREAYEDYRHFSSCEPPIPREAGEAYDFIPTSMDFPQORFRALANQVVG 120
DB 61 GHWIATRGQILREAYEDYRHFSSCEPPIPREAGEAYDFIPTSMDFPQORFRALANQVVG 120
QY 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYL 180
DB 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYL 180
QY 181 TDQWTRPDGSMTEFAEAKAALYDYLPIIEQRORKEGTDAISIVANGQVNGRPITSDAKR 240

```

181 TDQMTREDSMTFAEAKENALDYLLPIIEQRQKPGTDAISIVNGQVNGRPITSDEAKR 240
241 MCGLLLVGGLETVVNFSLFSMEFLAKSPHQRQELIERPERIPAAACBELLRRFSLVADGRI 300
241 MCGLLLVGGLETVVNFSLFSMEFLAKSPHQRQELIERPERIPAAACBELLRRFSLVADGRI 300
301 LTSDFEHGVLKKGQDQILLPQMLSGLDERKNACPMHVDPSRQKVSHTTFHGSHLCLGQ 360
301 LTSDFEHGVLKKGQDQILLPQMLSGLDERKNACPMHVDPSRQKVSHTTFHGSHLCLGQ 360
361 HJARRELIWTKWLTRIPDFSAPGAQIOHKSGIVSGVQALPLVMDPATTKAV 414
361 HJARRELIWTKWLTRIPDFSAPGAQIOHKSGIVSGVQALPLVMDPATTKAV 414

RESULT 4
US-10-453-104-13
; Sequence 13, Application US/10453104
; Publication No. US20030207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Hyun Joo
; TITLE OF INVENTION: Oxxygenase Enzymes and Screening Method
; FILE REFERENCE: 4058/1827-US3
; CURRENT APPLICATION NUMBER: US/10/453,104
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 09/661,093
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 09/246,451
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/394,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/686,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/106,834
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant M7-8H

US-10-453-104-13
Query Match 99.5%; Score 2169; DB 15; Length 414;
Best Local Similarity 99.8%; Pred. No. 9.3e-211;
Matches 413; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTETIQSANLAPPPHVEHLVDFDMYNLSAGVQEAWAVLQESNVPLVWTRCNG 60
DB 1 TTETIQSANLAPPPHVEHLVDFDMYNLSAGVQEAWAVLQESNVPLVWTRCNG 60
QY 61 GHWIATRGQIREAYEDYRHSSECFPIPRAGEAYDFTPTSDPPEQORFALANQVVG 120
DB 61 GHWIATRGQIREAYEDYRHSSECFPIPRAGEAYDFTPTSDPPEQORFALANQVVG 120
QY 121 MPVVDKLENRQELACSLIESLRPOGQNFEDYAEPPPIRIFMLLAGLPEEDIPLHXYL 180
DB 121 MPVVDKLENRQELACSLIESLRPOGQNFEDYAEPPPIRIFMLLAGLPEEDIPLHXYL 180
QY 181 TDQMTREDSMTFAEAKENALDYLLPIIEQRQKPGTDAISIVNGQVNGRPITSDEAKR 240
DB 181 TDQMTREDSMTFAEAKENALDYLLPIIEQRQKPGTDAISIVNGQVNGRPITSDEAKR 240
QY 241 MCGLLLVGGLETVVNFSLFSMEFLAKSPHQRQELIERPERIPAAACBELLRRFSLVADGRI 300
DB 241 MCGLLLVGGLETVVNFSLFSMEFLAKSPHQRQELIERPERIPAAACBELLRRFSLVADGRI 300

301 LTSDFEHGVLKKGQDQILLPQMLSGLDERKNACPMHVDPSRQKVSHTTFHGSHLCLGQ 360
301 LTSDFEHGVLKKGQDQILLPQMLSGLDERKNACPMHVDPSRQKVSHTTFHGSHLCLGQ 360
361 HJARRELIWTKWLTRIPDFSAPGAQIOHKSGIVSGVQALPLVMDPATTKAV 414
361 HJARRELIWTKWLTRIPDFSAPGAQIOHKSGIVSGVQALPLVMDPATTKAV 414

RESULT 5
US-10-214-446-50
; Sequence 50, Application US/10214446
; Publication No. US20030180742A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, David
; APPLICANT: Burk, Mark J.
; APPLICANT: Hitchman, Tim
; APPLICANT: Fujol, Catherine
; APPLICANT: Richardson, Toby
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: P450 ENZYMES, NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: THEM AND METHODS OF MAKING AND USING THEM
; FILE REFERENCE: 09010-5003C1
; CURRENT APPLICATION NUMBER: US/10/214,446
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 60/309,497
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Bacterial
US-10-214-446-50
Query Match 17.3%; Score 377.5; DB 14; Length 404;
Best Local Similarity 27.6%; Pred. No. 3.6e-29;
Matches 97; Conservative 64; Mismatches 157; Indels 33; Gaps 5;

QY 60 GHWIATRGQIREAYEDYRHSSECFPIPRAGEAYDFTPTSDPPEQORFALANQVVG 103
DB 32 GHWIATRGQIREAYEDYRHSSECFPIPRAGEAYDFTPTSDPPEQORFALANQVVG 85
QY 104 --DPPEQORFALANQVVGMPVVDKLENRQELACSLIESLRPOGQNFEDYAEPPPIR 161
DB 86 LHDQPEHQLRQMQGFTTRLLITTEPKIORVCELLIDAFVKRGSTEFMTYAHPEPAK 145
QY 162 IFMLLAGLPEEDIPLHXYLTDQMTREDSMTFAEAKENALDYLLPIIEQRQ 213
DB 146 VIAEMGVNPDYPAFVVMSEDLNFAFAGSLRPTLEMFRAAQDGLLAWMDYFARLLPERE 205
QY 214 KPGTDAISIVANGQVNGRPITSDEAKRMCGLLLVGGLETVVNFSLFSMEFLAKSPHQR 273
DB 206 NEGDDLVLILLSAEGEGWMTAEQVLANTQIIVAGHETTRNLVANGVELLRYPEQAL 265
QY 274 LTERPERIPAAACBELLRRFS-LVADGRILTSYEHGVLKKGQDQILLPQMLSGLDERKN 332
DB 266 LESRPELMPSAVRMEIRFESPLQFIRRVAREDFEFGAEVRREGDGLVLMGSLANRDEAF 325
QY 333 ACPMHVDPSRQKVSHTTFHGSHLCLGQELARREIIVTLKEMLTRIPDFS 383
DB 326 DDPTDFDLTRNPTGELAFGWGPHVCVGAALAELEGQVSPFRLILDLRUGLE 376

RESULT 6
US-10-214-446-40
; Sequence 40, Application US/10214446
; Publication No. US20030180742A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, David
; APPLICANT: Burk, Mark J.
; APPLICANT: Hitchman, Tim
; APPLICANT: Fujol, Catherine


```

; APPLICANT: Richardson, Toby
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: P450 ENZYMES, NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: THEM AND METHODS OF MAKING AND USING THEM
; FILE REFERENCE: 09010-5000C1
; CURRENT APPLICATION NUMBER: US/10/214,446
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 60/309,497
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Bacterial
US-10-214-446-40

Query Match      17.3%; Score 376.5; DB 14; Length 404;
Best Local Similarity 30.1%; Pred. No. 4.6e-29;
Matches 119; Conservative 58; Mismatches 183; Indels 31; Gaps 10;

QY 31 PNLSAGVQEAUWVQESNVDLVWTRCNGGH--WIATRGQILREAYEDYRHFSSECPFI 88
Db 20 PENEADGISLADAYEEAREQPGGLLRVMAYGEPALATR-----YADARLVLDGRFS 72
QY 89 -----PREAGEAYDPIPTSMDEPQEQFRALANOVVGMPPVDKLENRIQELACSLI 139
Db 73 RAEGARHDEPRQEGEDSGILSDPDHTRLRITLVAKFTMKEVQKLRPAVRELADLI 132
QY 140 ESLRPGQ-QCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYLTDQWTRPDGSMW---FAE 195
Db 133 DKMWATGAPVDLVEEFAIPVGVICQLLGPVEDRPRFRAWSD-AALSTSSITABEFDA 191
QY 196 AKCALDYLLPIITEQRQKPGTDAISIVANGQVNGRPITSDEAKRMCGLLVGLQTVVN 255
Db 192 NOELRAYMRGLIEDHRAEPREDLTGLIEARDORDRLTEQELVDCVGLVAGHETTAT 251
QY 256 FLFSMEFLAKSPHQRQELTERPERIPAAACEELLRRFSLVADG---RLTSDYEFHGVQ 311
Db 252 QINFTVUTLLDRQEWNRNLRDEDELVPTAVEELM-REVPLGSSMSFRYATEDVEVGTL 310
QY 312 LKXGDQILLPQML-SGLDERKXNACPMHVDPSRQKVSHTTGHGSHLCUGQHLAREIIVTL 371
Db 311 VRAGEPVLVAVGAANRDPARFADPAQELDLAREGNCILGFGHGHVHCLGAPLARLEQAL 370
QY 372 KEMLTIPDSIAPGQIQHKS-GIVSGVQALPLVW 406
Db 371 GALLRLPLGRIIA--GDIEWKTMVLVGRPTLPVGM 404

```

```

RESULT 7
US-10-314-657-37
; Sequence 37, Application US/10314657
; Publication No. US20030175888A1
; GENERAL INFORMATION:
; APPLICANT: SHEN, Ben
; APPLICANT: CHENG, Yi-Qiang
; APPLICANT: TANG, Gong-Li
; TITLE OF INVENTION: Discrete Acyltransferases Associated with Type I Polyketide
; TITLE OF INVENTION: Syntheses and Methods of Use
; FILE REFERENCE: 034030-0021
; CURRENT APPLICATION NUMBER: US/10/314,657
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US02/08937
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 60/278,935
; PRIOR FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIt version 3.2
; SEQ ID NO 37
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Streptomyces atroolivaceus

```

US-10-314-657-37

```

Query Match      17.1%; Score 372; DB 14; Length 399;
Best Local Similarity 28.1%; Pred. No. 1.3e-28;
Matches 110; Conservative 70; Mismatches 176; Indels 36; Gaps 9;

QY 38 VOEAWAVQESNVDLVWTRCNGGHWIATRGQILREAYEDYRHFSSECPFIPREAGE---94
Db 21 IHPKFAELRETOPLARVRLPYGEGMMVTR-----YDDVRAANSDFRFSRAQIGEDTFF 73
QY 95 -----AYDFIPTSMDPEQEQFRALANOVVGMPPVDKLENRIQELACSLIESLRPQOQ 147
Db 74 RTTELAARSDTI--LSLDPEPHTKRLRLSKAFTARRRGAMQSWLEELFAGLLDGVERTGH 132
QY 148 -CNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYLTDQWTRPDGSMTPAEAKE-----A 199
Db 133 PADIVEDDLAQPTTIATVICRLGVPYEDRGRFQHWSEVI-----MSTAYSKEEAVSADAS 187
QY 200 LVDYLPIITEQRQKPGTDAISIVANGQVNGRPITSDEAKRMCGLLVGLDVTVNFSLF 259
Db 188 IRAYLADLVASARRAAPDDDLGLVLSARDDDRLTEDELITPGVTLLVAGHETSAHQJGN 247
QY 260 SMEFLAKSPHQRQELTERPERIPAAACEELLRRFSL---VADGRILTSDFPHGVQLKKGD 316
Db 248 MVYALLTHEDQLSLLREQPELLPRAVEELLRFPVLGNGVGNARIALDEVELSGGTVEAGE 307
QY 317 QILLPQML-SGLDERKXNACPMHVDPSRQKVSHTTTFHGHSHLCUGQHLARRIIVTLKEMLT 376
Db 308 GVVAANVANRDPRAFDDPDLRLDITREKNPHLAFHGHAHYCLGQALARMELRVAIGGLE 367
QY 377 RIPDFSIA-PGAIQIHK-S-GIVSGVQALPLVW 406
Db 368 RFEGSLAVPADQVWKTGGLFRGPQRLPIAW 399

```

RESULT 8

```

US-10-156-761-14997
; Sequence 14997, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 151c9
; SEQ ID NO 14997
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14997

```

```

Query Match      16.8%; Score 367; DB 14; Length 416;
Best Local Similarity 29.1%; Pred. No. 4.4e-28;
Matches 104; Conservative 56; Mismatches 181; Indels 16; Gaps 8;

QY 63 WIATRGQILREAYEDYRHFSSECPFIPR--EAGRAYDIPT-----SMDPPEQQRAL 114
Db 63 WVTGHAARALLSDQRLSSDRT--LPFRPATTTFEAVTRRVRVALLGVDDPPEHRTQRM 120
QY 115 ANCVGMPPVDKLENRIQELACSLIESLRPQ-QCNFTEDYAEPPPIRIFMLLAGLPEED 173
Db 121 LVPSFTLKRAAALRERIGETVDGLDAMEAQPPAELVSAPALPLFSVMICALLGVPPAD 180

```

Patent No. US20020164742A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600,438US1
CURRENT APPLICATION NUMBER: US/09/860,846
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 39
LENGTH: 416
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-860-846-39

Query Match 16.8%; Score 365.5; DB 9; Length 416;
Best Local Similarity 28.3%; Pred. No. 6.2e-28;
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

QY 76 EDYRHSSECPPIPREAGEAYDFTPSMDPPQRCQFRALANQVGMVVDKLENRIQELA 135
DB 72 KDWK--NSTPLTEAEALNHNMLE--DPRHTRLRKLVAEFTMRRVELLRPRVQEI 127
QY 136 CSLIESL--RPOQCNTEDYAEPPPIRIFMILLAGLPEEDIPLHKYLTQMTPEPGSMTF 193
DB 128 DGLVDAMLAAPDGRADLMESLAWPLFITVISELLGVPEPDRAAFRVWTDVFPDDPAQA 187
QY 194 AEAKALYDYLPIEQRQKPGTDAIS-IVANGQVNGRPITSDEAKRMCGLLLVGGLDT 252
DB 188 QTAMAMSGYLSRLDSKGGQDGLLSALVRTSDEGSLTSEELGMAHILLVAGHET 247
QY 253 VYNFLSFSMEFLAKSPEHQELIERPERIPAAACELLRFRSLVADGRILTSDFE 307
DB 248 TVNLIANGMYALLSHPDQALRADMTLLDGAVEEMLR-----YEGVESATYRFPV 302
QY 308 --HGVLKKGQDILLPQMSGLDERKNACPMHVDFRQKVSHTTFGSHLCLGQHLARR 365
DB 303 DLDGTVIPAGDTVLVLAADAHRTPEFPDPHRRDIRDTAGHLAFGHGHCIGAPLARL 362
QY 366 EIVTLKWLTRIPDFS--IAPGAQIQHKSGLVSGVQALPLVW 406
DB 363 EARIAVALLERCPDIALDVSPGELVWYPMIRGLKALPIRW 405

RESULT 11
US-09-988-384B-39
Sequence 39, Application US/09988384B
Publication No. US20030073824A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600,536US1
CURRENT APPLICATION NUMBER: US/09/988,384B
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: PCT/US99/14398
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 53
SEQ ID NO 39
LENGTH: 416
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-988-384B-39

174 IPEHLKYLTDQMTPEPGSMTFPAEAKALYDYLPIEQRQKPGTDAISIVANGQVNGRPI 233
DB 181 HDEFESQRELLRGPAGAEVQDARAQDDVLYALIDRKRKEPGDGLDDLTQEQINRGTV 240
QY 234 TSEAKRMCGLLLVGGLDTVNVLSFMSFELAKSPEHQELIERPERIPACBELLRRFS 293
DB 242 DRAELVSLATLLLAGHETANMISLGTFTLLRHPQELAEARAGPGLMPAAVEELL-RFL 299
QY 294 LVADG--RILTSDFEFGVOLKGGDQILLPQMSGLDERKNACPMHVDPSQKVSHTTFG 351
DB 300 SIADGLLEVATEDIEVAGTIRADGGVENSVINRDAAGFAEPDADLWHRSARHHVAFG 359
QY 352 HSHLCLGQHLAREIIVTLKWTTRIPDFS-IAPGAQIQHKSGLVSGVQALPLVW 406
DB 360 FGIQCLGQNLARAEMETALGTLERLPGLRAAPADIPFPKSDTIQGMLELFTV 416

RESULT 9
US-09-861-289-39
Sequence 39, Application US/09861289
Patent No. US20020110897A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600,438US1
CURRENT APPLICATION NUMBER: US/09/861,289
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 39
LENGTH: 416
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-861-289-39

Query Match 16.8%; Score 365.5; DB 9; Length 416;
Best Local Similarity 28.3%; Pred. No. 6.2e-28;
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

QY 76 EDYRHSSECPPIPREAGEAYDFTPSMDPPQRCQFRALANQVGMVVDKLENRIQELA 135
DB 72 KDWK--NSTPLTEAEALNHNMLE--DPRHTRLRKLVAEFTMRRVELLRPRVQEI 127
QY 136 CSLIESL--RPOQCNTEDYAEPPPIRIFMILLAGLPEEDIPLHKYLTQMTPEPGSMTF 193
DB 128 DGLVDAMLAAPDGRADLMESLAWPLFITVISELLGVPEPDRAAFRVWTDVFPDDPAQA 187
QY 194 AEAKALYDYLPIEQRQKPGTDAIS-IVANGQVNGRPITSDEAKRMCGLLLVGGLDT 252
DB 188 QTAMAMSGYLSRLDSKGGQDGLLSALVRTSDEGSLTSEELGMAHILLVAGHET 247
QY 253 VYNFLSFSMEFLAKSPEHQELIERPERIPAAACELLRFRSLVADGRILTSDFE 307
DB 248 TVNLIANGMYALLSHPDQALRADMTLLDGAVEEMLR-----YEGVESATYRFPV 302
QY 308 --HGVLKKGQDILLPQMSGLDERKNACPMHVDFRQKVSHTTFGSHLCLGQHLARR 365
DB 303 DLDGTVIPAGDTVLVLAADAHRTPEFPDPHRRDIRDTAGHLAFGHGHCIGAPLARL 362
QY 366 EIVTLKWLTRIPDFS--IAPGAQIQHKSGLVSGVQALPLVW 406
DB 363 EARIAVALLERCPDIALDVSPGELVWYPMIRGLKALPIRW 405

RESULT 10
US-09-860-846-39
Sequence 39, Application US/09860846

Query Match 16.8%; Score 365.5; DB 10; Length 416;
Best Local Similarity 28.3%; Pred. No. 6.2e-28;
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

QY 76 EDYRHSSECPPIPREAGEAYDFTPTSDPPEQRFALANQVGMVVDKLENRIQELA 135
DB 72 KQWR--NSTTPTLTERAALNNHML--DPPHTRKRLVAREFTMRVRLPRVQEV 127

QY 136 CSLIESL--RPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYLTDQMTDPGSMTF 193
DB 128 DGLVDMMLAAPDGRADLMESLAWPLITVISSELLGVPEPDRAAFRVWIDAFVFPDPAQA 187

QY 194 ABAEALYDYLIPIIEQRQKPGTDAIS--IVANGVNGRPITSDAKMCGLLVGGD 252
DB 188 QTAMAEMSGYLSRLIDSKRGQDGEDLLSALVTSDEDSRLTSEELLGMAHILLVAGHET 247

QY 253 VVNFSLFSMEFLAKSPEHRQELIERPERIPACEELLRRFSLVADGRILTSYEF---- 307
DB 248 TVNLITANGMYALLSHPDQALRADMTLLDGAVEMLR-----YEGPVESATYRFPV 302

QY 308 --HGVLKKGQDQILLPQMLSGLDERKNACPMHVDFSROKVSHTTTFHGSHLCLGQHLARR 365
DB 303 DLDTGTVIPAGDVLAVLADAHRTPERFDPHFRDTRDTAGHLAFGHGTHFCIGAPLARI 362

QY 366 EIIVTLKEMWLRIPDFS--IAPGAQIQHKSGLVSGVQALPLVW 406
DB 363 EARIARALLERCPLDALDVSPGELVWYPMIRGLKALPIRW 405

RESULT 13
US-09-793-708-18
; Sequence 18, Application US/09793708
; Publication No. US20030104597A1
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONCLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002121
; CURRENT APPLICATION NUMBER: US/09/793,708
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 09/657,440
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 09/320,878
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: US 09/141,908
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: US 09/073,538
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: US 08/846,247
; PRIOR FILING DATE: 1997-04-30
; PRIOR APPLICATION NUMBER: US 60/134,990
; PRIOR FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patent in ver. 2.0
; SEQ ID NO 18
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-793-708-18

Query Match 16.8%; Score 365.5; DB 10; Length 416;
Best Local Similarity 28.3%; Pred. No. 6.2e-28;
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

QY 76 EDYRHSSECPPIPREAGEAYDFTPTSDPPEQRFALANQVGMVVDKLENRIQELA 135
DB 72 KQWR--NSTTPTLTERAALNNHML--DPPHTRKRLVAREFTMRVRLPRVQEV 127

QY 136 CSLIESL--RPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYLTDQMTDPGSMTF 193
DB 128 DGLVDMMLAAPDGRADLMESLAWPLITVISSELLGVPEPDRAAFRVWIDAFVFPDPAQA 187

QY 194 ABAEALYDYLIPIIEQRQKPGTDAIS--IVANGVNGRPITSDAKMCGLLVGGD 252
DB 188 QTAMAEMSGYLSRLIDSKRGQDGEDLLSALVTSDEDSRLTSEELLGMAHILLVAGHET 247

QY 253 VVNFSLFSMEFLAKSPEHRQELIERPERIPACEELLRRFSLVADGRILTSYEF---- 307
DB 248 TVNLITANGMYALLSHPDQALRADMTLLDGAVEMLR-----YEGPVESATYRFPV 302

QY 308 --HGVLKKGQDQILLPQMLSGLDERKNACPMHVDFSROKVSHTTTFHGSHLCLGQHLARR 365
DB 303 DLDTGTVIPAGDVLAVLADAHRTPERFDPHFRDTRDTAGHLAFGHGTHFCIGAPLARI 362

QY 366 EIIVTLKEMWLRIPDFS--IAPGAQIQHKSGLVSGVQALPLVW 406
DB 363 EARIARALLERCPLDALDVSPGELVWYPMIRGLKALPIRW 405

RESULT 14
US-10-201-365-13

Query Match 16.8%; Score 365.5; DB 10; Length 416;
Best Local Similarity 28.3%; Pred. No. 6.2e-28;
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

QY 76 EDYRHSSECPPIPREAGEAYDFTPTSDPPEQRFALANQVGMVVDKLENRIQELA 135
DB 72 KQWR--NSTTPTLTERAALNNHML--DPPHTRKRLVAREFTMRVRLPRVQEV 127

QY 136 CSLIESL--RPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYLTDQMTDPGSMTF 193
DB 128 DGLVDMMLAAPDGRADLMESLAWPLITVISSELLGVPEPDRAAFRVWIDAFVFPDPAQA 187

QY 194 ABAEALYDYLIPIIEQRQKPGTDAIS--IVANGVNGRPITSDAKMCGLLVGGD 252
DB 188 QTAMAEMSGYLSRLIDSKRGQDGEDLLSALVTSDEDSRLTSEELLGMAHILLVAGHET 247

QY 253 VVNFSLFSMEFLAKSPEHRQELIERPERIPACEELLRRFSLVADGRILTSYEF---- 307
DB 248 TVNLITANGMYALLSHPDQALRADMTLLDGAVEMLR-----YEGPVESATYRFPV 302

QY 308 --HGVLKKGQDQILLPQMLSGLDERKNACPMHVDFSROKVSHTTTFHGSHLCLGQHLARR 365
DB 303 DLDTGTVIPAGDVLAVLADAHRTPERFDPHFRDTRDTAGHLAFGHGTHFCIGAPLARI 362

QY 366 EIIVTLKEMWLRIPDFS--IAPGAQIQHKSGLVSGVQALPLVW 406
DB 363 EARIARALLERCPLDALDVSPGELVWYPMIRGLKALPIRW 405

RESULT 12
US-09-836-821-39
; Sequence 39, Application US/09836821
; Publication No. US20030037405A1
; GENERAL INFORMATION:
; APPLICANT: Shermat, D. H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, Z.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.43801
; CURRENT APPLICATION NUMBER: US/09/836,821
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-836-821-39

Query Match 16.8%; Score 365.5; DB 10; Length 416;
Best Local Similarity 28.3%; Pred. No. 6.2e-28;
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

QY 76 EDYRHSSECPPIPREAGEAYDFTPTSDPPEQRFALANQVGMVVDKLENRIQELA 135
DB 72 KQWR--NSTTPTLTERAALNNHML--DPPHTRKRLVAREFTMRVRLPRVQEV 127

QY 136 CSLIESL--RPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYLTDQMTDPGSMTF 193
DB 128 DGLVDMMLAAPDGRADLMESLAWPLITVISSELLGVPEPDRAAFRVWIDAFVFPDPAQA 187

QY 194 ABAEALYDYLIPIIEQRQKPGTDAIS--IVANGVNGRPITSDAKMCGLLVGGD 252
DB 188 QTAMAEMSGYLSRLIDSKRGQDGEDLLSALVTSDEDSRLTSEELLGMAHILLVAGHET 247

QY 253 VVNFSLFSMEFLAKSPEHRQELIERPERIPACEELLRRFSLVADGRILTSYEF---- 307
DB 248 TVNLITANGMYALLSHPDQALRADMTLLDGAVEMLR-----YEGPVESATYRFPV 302

QY 308 --HGVLKKGQDQILLPQMLSGLDERKNACPMHVDFSROKVSHTTTFHGSHLCLGQHLARR 365

Sequence 13, Application US/10201365
Publication No. US20030148469A1
GENERAL INFORMATION:
APPLICANT: ASHLEY, Gary
APPLICANT: BETLACH, Melanie C.
APPLICANT: BETLACH, Mary
APPLICANT: MCDANIEL, Robert
APPLICANT: TANG, Li
TITLE OF INVENTION: COMBINATORIAL POLYKETIDE LIBRARIES PRODUCED USING A MODULAR
TITLE OF INVENTION: PKS GENE CLUSTER AS SCAFFOLD
FILE REFERENCE: 300622002103
CURRENT APPLICATION NUMBER: US/10/201,365
CURRENT FILING DATE: 2002-07-22
PRIOR APPLICATION NUMBER: US 09/14,908
PRIOR FILING DATE: 1998-08-28
PRIOR APPLICATION NUMBER: US 09/073,538
PRIOR FILING DATE: 1998-05-06
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patent in Ver. 2.0
SEQ ID NO 13
LENGTH: 416
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-10-201-365-13

Query Match 16.8%; Score 365.5; DB 14; Length 416;
Best Local Similarity 28.3%; Pred. No. 6.2e-28;
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

QY 76 EDYRHFSECCFIPREAGEAYDFIPTSDPPQRCFRALANQVGMVVPVVDKLENRIQELA 135
Db 72 KQWR--NSTTPTLTERAEALNNHLES--DPPHTLRLKLVAREFTMRRAVELLRPRVQELV 127
QY 136 CSLIESL--RPGQCQNFTEVAEPPPIRIFMLLAGLPEDIPHLKYLTDQMTRPDGSMTF 193
Db 128 DGLVDMLAAPDGRADLMESLAWPLTIVISELLGVPEPDRAAFRVMTDAFVFPDDPAQA 187
QY 194 ABAKEALVDYLPIITEQRQKPGTDAIS-IVANGQVNGRPITSDAKRMCGLLLVGGLDT 252
Db 188 QTAAEMSGYLSRLIDSKRGQDGEDLLSALVTSDEDSRLTSEELGMAHILLVAGHET 247
QY 253 VVNFLSFMEFLAKSPHEHQELIERPERIPAAACEILLRFRSLVADGRILTSDYEF----- 307
Db 248 TVNLIANGMYALLSHFDQLAALRADMTLLDGAVEMLR-----YEGPVESATYRFPVEPV 302
QY 308 --HGVLKKGQILLPQMLSGLDERKNACPMFVDFSRQKVSHTTGHSHLCLGQHILARR 365
Db 303 DLDGTVIPAGDTLVVLADAHRTPERFDPHFRDIRDTAGHLAFGHGIFHCIGAPLARI 362
QY 366 ELIVTLKZWLTIPDFS--IAPGAIQHKSGIVSGVQALPLVW 406
Db 363 EARAVRALERCPLDALDVSEGLVWVWPNMIRGLKALPIRW 405

Search completed: April 6, 2004, 19:14:30
JOB time : 35.974 secs

RESULT 15
US-10-160-539-18
Sequence 18, Application US/10160539
Publication No. US20030162262A1
GENERAL INFORMATION:
APPLICANT: ASHLEY, Gary
APPLICANT: BETLACH, Melanie C.
APPLICANT: BETLACH, Mary C.
APPLICANT: MCDANIEL, Robert
APPLICANT: TANG, Li
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 300622002120
CURRENT APPLICATION NUMBER: US/10/160,539
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: US/09/657,440
PRIOR FILING DATE: 2000-09-07
PRIOR APPLICATION NUMBER: 09/320,878
PRIOR FILING DATE: 1999-05-27
PRIOR APPLICATION NUMBER: CIP OF 09/141,908

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 18:49:16 ; Search time 14.9588 Seconds
(without alignments)
1428.803 Million cell updates/sec

Title: US-09-246-451A-11

Perfect score: 2180

Sequence: 1 TTTETQSNANLAPLPPHPE.....IVSGVQALPLVWDPATKAV 414

Scoring table: BLOSUM62

Gapop 13.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/2/iaa/5A.COMB.pep.*
- 2: /cgn2_6/prodata/2/iaa/5B.COMB.pep.*
- 3: /cgn2_6/prodata/2/iaa/6A.COMB.pep.*
- 4: /cgn2_6/prodata/2/iaa/6B.COMB.pep.*
- 5: /cgn2_6/prodata/2/iaa/6C.COMB.pep.*
- 6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	365.5	16.8	416	3	US-09-320-878-18
2	365.5	16.8	416	3	US-09-105-537-39
3	365.5	16.8	416	4	US-09-141-908-13
4	365.5	16.8	416	4	US-09-657-440-18
5	349	16.0	437	4	US-09-252-991A-17836
6	340	15.6	403	6	5212296-5
7	333	15.3	409	3	US-09-385-028-12
8	333	15.3	409	4	US-09-726-614-12
9	333	15.3	409	4	US-09-385-040-12
10	330	15.1	406	6	5212296-6
11	324.5	14.9	399	4	US-08-765-907A-10
12	316	14.5	412	1	US-08-102-863-11
13	316	14.5	412	5	PCT-US92-10885-11
14	271.5	12.5	419	3	US-09-335-409-8
15	271.5	12.5	419	3	US-09-413-814-71
16	271.5	12.5	419	4	US-09-568-102-8
17	271.5	12.5	419	4	US-09-567-969-8
18	271.5	12.5	419	4	US-09-568-480-8
19	271.5	12.5	419	4	US-09-568-486-8
20	271.5	12.5	419	4	US-09-568-472-8
21	271.5	12.5	419	4	US-09-567-899-8
22	261	12.0	395	4	US-09-262-965-129
23	234	10.7	468	4	US-09-252-991A-32437
24	168.5	7.7	189	4	US-09-679-279-20
25	160	7.3	422	2	US-09-096-982-5
26	160	7.3	422	2	US-08-653-650A-5
27	160	7.3	474	2	US-09-096-982-8

28	160	7.3	474	2	US-08-653-650A-8	Sequence 8, Appli
29	157	7.2	443	2	US-09-096-982-9	Sequence 9, Appli
30	157	7.2	443	2	US-08-653-650A-9	Sequence 9, Appli
31	155	7.1	422	1	US-08-396-218-2	Sequence 2, Appli
32	155	7.1	422	1	US-08-760-116-2	Sequence 2, Appli
33	130	6.0	512	2	US-08-194-981B-5	Sequence 5, Appli
34	127	5.8	516	4	US-09-215-694-16	Sequence 16, Appli
35	123.5	5.7	382	3	US-09-320-878-7	Sequence 7, Appli
36	123.5	5.7	382	4	US-09-141-908-7	Sequence 7, Appli
37	123.5	5.7	382	4	US-09-657-440-7	Sequence 22, Appli
38	123.5	5.7	402	3	US-09-105-537-22	Sequence 2, Appli
39	123.5	5.7	503	4	US-09-583-447A-2	Sequence 4, Appli
40	123.5	5.7	3782	3	US-09-105-537-4	Sequence 4, Appli
41	123	5.6	504	4	US-09-583-447A-4	Sequence 4, Appli
42	122	5.6	503	4	US-09-144-167-2	Sequence 24, Appli
43	121.5	5.6	524	4	US-09-126-420A-24	Sequence 6, Appli
44	118	5.4	513	3	US-08-948-564-6	Sequence 3, Appli
45	113.5	5.2	490	1	US-08-201-118-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-320-878-18
; Sequence 18, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320,878A
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119,139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100,880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-320-878-18

Query Match	16.8%	Score	365.5	DB 3	Length	416
Best Local Similarity	28.3%	Pred. No.	2.1e-29			
Matches	97	Conservative	62	Mismatches	163	Gaps 7
Qy	76	EDVHPSGCPFIPREAGEAYDFIPTSDPPEQRQFRALANQVVGMPYVDKLENNRIQBLA	135			
Db	72	KDWR--NSTPLTAEALNNHML--DPRHTRLKLVAAREFTMRVRLPRVQIV	127			
Qy	136	CSLIESL--RPQGCNFTEDYAEFPFIRIPLMLAGLPEEDIPHLKYLTDQMTDQSGMTFF	193			
Db	128	DGLVDMALAAPDGRADLMESLAWPLTITVISSELLGVPEPDRAAFRVMTDAFVFPDPAQA	187			
Qy	194	AEAKEALYDLIPIEIQROKPGTDAIS--IVANQVNGRPITSDAEKMGCELLLVGGSDT	252			
Db	188	QTAMAENSGYLSK--LDSKRGQDGEDLLSALVRTSDEDSRLTSBELLMGAILLVAGHET	247			

```

QY 253 VNFSLFSMEFLAKSPHROELIERPERIPAAACEBLLRRFSLVADGRILTSDYEF----- 307
Db 248 TVNLIANGMYALLSHPDQALRADMT-LDGAVEMLR-----YEGVESATYRFPPEPV 302
QY 308 --HGVCLKKGQDILLPQMLSGDLERKKNACPMHVDGFRQKVSHTTTFHGSHLCLGQHARR 365
Db 303 DLDTGTVIPAGDTVLVLADAHRTPERPDHREDIRDTAGHLAFGHHGHCIGAPLARR 362
QY 366 EIIIVTKENWTRIPDFS--IAPGAQIOHKSIVSGVQALPLVW 406
Db 363 EARIARALLERCPLDALDVSFGBLVWYFENPMIRGLKALPIRW 405

```

RESULT 2

```

US-09-105-537-39
; Sequence 39, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-105-537-39

```

Query Match 16.8%; Score 365.5; DB 3; Length 416;

Best Local Similarity 28.3%; Pred. No. 2.1e-29;

Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

```

QY 76 EDYRHSSECFPTPREAGEAYDE-PTSMDBPEQRFALANQVGMVVDKLENRIQELA 135
Db 72 KDWK--NSTTTLTEAEALNHNMLE--DPPRHTLRKLVAEFTMRVRLPRVQEI 127
QY 136 CSLTESL--RPQGCNFTEDYAEPPFIRIFMLLAGLPEEDIPHLKYLTDQMTDPGSMTF 193
Db 128 DGLVDMLAAPDGRADLMESLAWPLITV-SELLGVPEPDAFRVWTDAPVFPDPAQA 187
QY 194 AEAKALYDILPIEORRQKPGTDAIS-IVANQVNGRPITSDEAKMCGLLLVGGLDT 252
Db 188 QTAWAEMSGVLSRLIDSKRGQDGEDLLSALVTSDEDSRLTSBELLAGMAHILLVAGHET 247
QY 253 VNFSLFSMEFLAKSPHROELIERPERIPAAACEBLLRRFSLVADGRILTSDYEF----- 307
Db 248 TVNLIANGMYALLSHPDQALRADMT-LDGAVEMLR-----YEGVESATYRFPPEPV 302
QY 308 --HGVCLKKGQDILLPQMLSGDLERKKNACPMHVDGFRQKVSHTTTFHGSHLCLGQHARR 365
Db 303 DLDTGTVIPAGDTVLVLADAHRTPERPDHREDIRDTAGHLAFGHHGHCIGAPLARR 362
QY 366 EIIIVTKENWTRIPDFS--IAPGAQIOHKSIVSGVQALPLVW 406
Db 363 EARIARALLERCPLDALDVSFGBLVWYFENPMIRGLKALPIRW 405

```

RESULT 3

```

US-09-141-908-13
; Sequence 13, Application US/09141908
; Patent No. 6503741
; GENERAL INFORMATION:
; APPLICANT: Ashley, Gary
; APPLICANT: Betlach, Melanie C.
; APPLICANT: Mcdaniel, Robert
; APPLICANT: Tang, Li

```

```

; TITLE OF INVENTION: Combinatorial Polyketide Libraries Produced Using a
; TITLE OF INVENTION: Modular PKS Gene Cluster as Scaffold
; FILE REFERENCE: 300622002100
; CURRENT APPLICATION NUMBER: US/09/141,908
; CURRENT FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: PROV. 60/076,919
; EARLIER FILING DATE: 1998-03-05
; EARLIER APPLICATION NUMBER: PROV. 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-141-908-13

```

Query Match 16.8%; Score 365.5; DB 4; Length 416;

Best Local Similarity 28.3%; Pred. No. 2.1e-29;

Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

```

QY 76 EDYRHSSECFPTPREAGEAYDE-PTSMDBPEQRFALANQVGMVVDKLENRIQELA 135
Db 72 KDWK--NSTTTLTEAEALNHNMLE--DPPRHTLRKLVAEFTMRVRLPRVQEI 127
QY 136 CSLTESL--RPQGCNFTEDYAEPPFIRIFMLLAGLPEEDIPHLKYLTDQMTDPGSMTF 193
Db 128 DGLVDMLAAPDGRADLMESLAWPLITV-SELLGVPEPDAFRVWTDAPVFPDPAQA 187
QY 194 AEAKALYDILPIEORRQKPGTDAIS-IVANQVNGRPITSDEAKMCGLLLVGGLDT 252
Db 188 QTAWAEMSGVLSRLIDSKRGQDGEDLLSALVTSDEDSRLTSBELLAGMAHILLVAGHET 247
QY 253 VNFSLFSMEFLAKSPHROELIERPERIPAAACEBLLRRFSLVADGRILTSDYEF----- 307
Db 248 TVNLIANGMYALLSHPDQALRADMT-LDGAVEMLR-----YEGVESATYRFPPEPV 302
QY 308 --HGVCLKKGQDILLPQMLSGDLERKKNACPMHVDGFRQKVSHTTTFHGSHLCLGQHARR 365
Db 303 DLDTGTVIPAGDTVLVLADAHRTPERPDHREDIRDTAGHLAFGHHGHCIGAPLARR 362
QY 366 EIIIVTKENWTRIPDFS--IAPGAQIOHKSIVSGVQALPLVW 406
Db 363 EARIARALLERCPLDALDVSFGBLVWYFENPMIRGLKALPIRW 405

```

RESULT 4

```

US-09-657-440-18
; Sequence 18, Application US/09657440
; Patent No. 6509455
; GENERAL INFORMATION:
; APPLICANT: Ashley, Gary
; APPLICANT: Betlach, Melanie C.
; APPLICANT: Betlach, Mary C.
; APPLICANT: Mcdaniel, Robert
; APPLICANT: Tang, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/657,440
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/320,878
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: CIP OF 09/141,908
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 416
; TYPE: PRT

```

GENERAL INFORMATION:

GENERAL INFORMATION:

Db 175 EDRSAVLIDRGTYE---QVAKARDELVDGVLRELVEREINENPGTDLISRLVIDQVRPHL 231
 QY 234 TSDEAKMCGLLLVGLDVTWNFLSFSMEFLAKSPEHQELIERPERIPAAACEILLRFS 293
 Db 232 RVEEMVPCRLLLVAGHGTTSQASLSLSLLTDPALAGRLTEDPALLPKAVEELLRFS 291
 QY 294 LVADG--RLTSDYEFHGVQKGGDQILL-PQMLSGLDERKNACPMHVDPSQKVSHTTFG 351
 Db 292 IVQNGLARAAVEDVOLDVLRAGEGVVLSLSAGNRDETVPDPRVDVDRDARHLAFG 351
 QY 352 HGSMLCLGQHLAR---REIIVTLKEMLTRIPDFSIA-PGAQIOHKSGIVS-GVQALPLVM 406
 Db 352 HGMECCGQWLARVLEILLAAVLRW---PGARLAVPFEELDFRHEVSSYGLGALPTVM 408
 QY 407 D 407
 Db 409 Z 409

RESULT 9
 US-09-385-040-12
 ; Sequence 12, Application US/09385040
 ; Patent No. 6589775
 ; GENERAL INFORMATION:
 ; APPLICANT: Jensen, Susan E
 ; APPLICANT: Aideo, Kwamena A
 ; APPLICANT: Paraskar, Ashish S
 ; TITLE OF INVENTION: DNA SEQUENCE ENCODING ENZYMES OF CLAVULANIC ACID
 ; FILE REFERENCE: 09/385,040
 ; CURRENT APPLICATION NUMBER: US/09/385,040
 ; PRIOR FILING DATE: 1999-08-30
 ; PRIOR APPLICATION NUMBER: US 08/790,462
 ; PRIOR FILING DATE: 1997-01-29
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 12
 ; LENGTH: 409
 ; TYPE: PRT
 ; ORGANISM: Streptomyces clavuligerus
 US-09-385-040-12

Query Match 15.3%; Score 333; DB 4; Length 409;
 Best Local Similarity 29.0%; Pred. No. 5.1e-26;
 Matches 122; Conservative 60; Mismatches 185; Indels 54; Gaps 18;
 QY 16 PHVPHLVDFEDMYPNLSAGVQBAWVLSQSNVPDLVWTRCNGH-WIATRGQLIREA 74
 Db 14 PAYPMHRVCPVD---PPQLAGLSQKAASRVT---LW---DGSQVWLVTSHAGARAV 62
 QY 75 YEDYRHFS-SECPFIP-----REAGEAYDFIPTSMDPPEQCFRA-----LANQV 118
 Db 63 LGDRFTAVTSAPGPFMLTRTSQVRANPESASF1--RMDDFQHSRLRMLTRDFLARA 120
 QY 119 VGM-PWVDKLENRIQELACSTESLRPQCNFTEDYAEPPFIRIFMLLAGPEEDIPHL 177
 Db 121 EALRAVREL---LDELGLLVKGERP---VDLVAGLTIPVPSRVITLFGAGDORREFI 174
 QY 178 K----YLTQWTRPGSMTFAKALYDYLIPIDQKQKGTDAISIVANGQVNGRPI 233
 Db 175 EDRSAVLIDRGTYE---QVAKARDELVDGVLRELVEREINENPGTDLISRLVIDQVRPHL 231
 QY 234 TSDEAKMCGLLLVGLDVTWNFLSFSMEFLAKSPEHQELIERPERIPAAACEILLRFS 293
 Db 232 RVEEMVPCRLLLVAGHGTTSQASLSLSLLTDPALAGRLTEDPALLPKAVEELLRFS 291
 QY 294 LVADG--RLTSDYEFHGVQKGGDQILL-PQMLSGLDERKNACPMHVDPSQKVSHTTFG 351
 Db 292 IVQNGLARAAVEDVOLDVLRAGEGVVLSLSAGNRDETVPDPRVDVDRDARHLAFG 351
 QY 352 HGSMLCLGQHLAR---REIIVTLKEMLTRIPDFSIA-PGAQIOHKSGIVS-GVQALPLVM 406
 Db 352 HGMECCGQWLARVLEILLAAVLRW---PGARLAVPFEELDFRHEVSSYGLGALPTVM 408

QY 407 D 407
 Db 409 Z 409

RESULT 10
 5212296-6
 ; Patent No. 5212296
 ; APPLICANT: DEAN, CAROLINE; HARDER, PATRICIA A.; LETO, KENNETH
 ; J.; O'KEEFE, DANIEL P.; OVER, CHARLES A.; ROMESSER, JAMES A.
 ; TEPPERMAN, JAMES M.
 ; TITLE OF INVENTION: EXPRESSION OF HERBICIDE METABOLIZING
 ; CYTOCHROMES
 ; NUMBER OF SEQUENCES: 19
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/569,781
 ; FILING DATE: 23-AUG-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 464,499
 ; FILING DATE: 12-JAN-1990
 ; APPLICATION NUMBER: 405,605
 ; FILING DATE: 11-SEP-1989
 ; SEQ ID NO: 6
 ; LENGTH: 406
 5212296-6

Query Match 15.1%; Score 330; DB 6; Length 406;
 Best Local Similarity 26.3%; Pred. No. 1e-25;
 Matches 104; Conservative 74; Mismatches 193; Indels 24; Gaps 12;
 QY 31 PSNLSAGVQ--EAWAVLQESNVPDLVWTRCNGH-WIATRGQLIREAYEDYR----- 79
 Db 17 PSNRSCPQLPDGYAQLRDTPCPLHRVTLYDGRQAWVTKEAARKLLGDPRLSSNRD 76
 QY 80 HFSSECPFIP--REAGEAYDFIPTSMDPPEQCFRANALANQVGMPPVDKLENRIQELACS 137
 Db 77 NFPAATSPRFAVRSPQAF---IGLDPEHGTRRRMTISEFTVKIKGMRPEVEEVVHG 132
 QY 138 LIESLRPQ--QCNTEDYAEPEFIRIFMLLAGPEEDIPHLKYLTDQMTPEQCSMTFAEA 196
 Db 133 FIDEMLAGPTADLVSQFALPVPSMVICRLGLVPYADHBFQDASKRLVQSTDRQSAITA 192
 QY 197 KEALYDYLIPITTEQRQKPGTDAL-SIVANGQVNGEPIITSDEAKRMCGLLLVGLDVTVN 255
 Db 193 RNDLAGYLDCLITQCTEFGAGLVGALVADQLANGE-IDREELISTAMULLIAGHETTAS 251
 QY 256 FLSFSMEFLAKSPEHQELIERPERIPAAACEILLRFSL--VADGRILTSDYEFHGVOLK 313
 Db 252 MTSLSVITLDDHPEQVAALRADRSIVPGAVEELLRYLAIDAGGRVATADIEVEGELIR 311
 QY 314 KEDQILLPQMLSGLDERKNACPMHVDPSQKVSHTTFGHGSHLCLGQHLARREIIVTKE 373
 Db 312 AGEGVIVVNSIANRGTGVYEDPDALDIHRSARHLLAFGFGVHOCGLQNLARLEVLNA 371
 QY 374 WLTRIPDFSIA-PGAQIOHKSG-IVSGVQALPLVM 406
 Db 372 IMDRVPTLAVPTEQLVLRPGTTIQGVNELPTVM 406

RESULT 11
 US-08-765-907A-10
 ; Sequence 10, Application JS/08765907A
 ; Patent No. 6352839
 ; GENERAL INFORMATION:
 ; APPLICANT: BLANC, Veronique
 ; APPLICANT: THIBAUT, Denis
 ; APPLICANT: BAWAS-JACQUES, Nathalie
 ; APPLICANT: BLANCHE, Francis
 ; APPLICANT: COUZET, Joel
 ; APPLICANT: BARRIERE, Jean-Claude
 ; APPLICANT: DEBUSSCHE, Laurent
 ; APPLICANT: FAMECHON, Alain

APPLICANT: PARIS, Jean-Marc
APPLICANT: DUTRUC-ROSSET, Gilles
TITLE OF INVENTION: Streptogramins And Method For Preparing Same By
TITLE OF INVENTION: Mutasynthesis
FILE REFERENCE: Streptogramin genes
CURRENT APPLICATION NUMBER: US/08/765,907A
CURRENT FILING DATE: 1997-03-20
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 12
LENGTH: 399
TYPE: PRT
ORGANISM: Streptomyces pristinaespiralis
US-08-765-907A-10

Query Match 14.9%; Score 324.5; DB 4; Length 399;
Best Local Similarity 29.0%; Pred. No. 3.8e-25;
Matches 106; Conservative 59; Mismatches 159; Indels 41; Gaps 12;

QY 74 AYEDYRH-----FSSECPFIPREAGEAYDPTSMDDPPEQKQFRALANQVGM 121
DB 36 AEFVRAADVLTVASDFGVSSQLRLRPGSQALSEQLLSWIDPPMHTLRLVSAFT 95
QY 122 PVVDKLENRQELACSLIESLRPGQC-NFTEDYAEFFIRIFMLLAGLPEEDIPHLKYL 180
DB 96 RTVADLEPRVTELAGQLDAV--DGDYFDLWADPAYPLPVIVIAELLGVPPADRTLFRSW 153
QY 181 TDQWTR-----PDGSMYFAEAKALYDYLPIIEQRQKPGTDAISIVA 224
DB 154 SDRMLQMQVADDMQGDGDADEYQRLVKEPMAHAYLHDVTDRRARPANDLISALV 213
QY 225 NGQVNGRPITSDAKRMCGLLVGLDVTWNFLSFSMEFLAKSPHQRQELIERPER--IP 282
DB 214 AARVEGRLEDEQIVERGALLMAGHVSSTMLGNVTVLCKDHP--RAEAAARADRSIIP 271
QY 283 AACBELLR-RFSLWADGRITTSYEFHGVQVKKGQILLQMLSL-GLDERKXNACPMHVD 340
DB 272 ALIEEVLELRPPITVMARVTTKDTVLAGTTIPAG-RMVVPSLLSANDEQVETDPDHL3L 330
QY 341 SRQKVSHTTFGHSHLCLGQHLARREIIVTLKWLTRIPDPSIAPGACIO-HKSGSIYGV 399
DB 331 AREG-RQIAFGHGIHYCLGAPLARLEGRALDEALDFRPFDPSTDGAKLRYHRDGLF-GV 388
QY 400 QALPL 404
DB 389 KNLPL 393

RESULT 12
US-08-102-863-11
Sequence 11, Application US/08102863
Patent No. 5466590
GENERAL INFORMATION:
APPLICANT: SARIASLANI, SIMA
TITLE OF INVENTION: CONSTITUTIVE
TITLE OF INVENTION: EXPRESSION OF P450SOY
TITLE OF INVENTION: AND FERRODIXIN-SOY IN
TITLE OF INVENTION: STREPTOMYCES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSEE: E. I. DU PONT DE NEMOURS
ADDRESSEE: AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: USA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0,
SOFTWARE: Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/102,863
FILING DATE: 435
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/807,001
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: GALLEGOS, R. THOMAS
REGISTRATION NUMBER: 32,692
REFERENCE/DOCKET NUMBER: CR-9000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-7342
TELEFAX: 302-892-7949
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 412 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-102-863-11

Query Match 14.5%; Score 316; DB 1; Length 412;
Best Local Similarity 27.2%; Pred. No. 3.1e-24;
Matches 84; Conservative 56; Mismatches 163; Indels 6; Gaps 5;
QY 103 MDPEQKQFRALANQVGMVVDKLENRQELACSLIESLRPGQ-QCNFTEDYAEFFIR 161
DB 105 VDDPEHTQRMILPTFSVKRIGALRPRIQTVDRLLDMERQGPAPBLVSAFALPVFSM 164
QY 162 IFMLLAGLPEEDIPHLKYLTDQWTRPDGSMYFAEAKALYDYLPIIEQRQKPGTDAIS 221
DB 165 VICALLGVVADHAPFEERSQRLRPGCADDVNRAROBLELYGALIDRKAEPDGLD 224
QY 222 IVANGVNGRPITSDAKRMCGLLVGLDVTWNFLSFSMEFLAKSPHQRQELIERPERI 281
DB 225 ELIHRDPDGPVDRQELVAFVITLLIAGHETTANMISLGTFTLLSHPEQLAALRAGTST 284
QY 282 PAACEELLRRFSLVADG--RILTSDYEFHGVQVKKGQILLQMLSGLDERKXNACPMHVD 339
DB 285 AVVEELL-RFLSIAEGQLRATEDMEVDGATIRKGVWFSTSLINRDAVFPRAETLD 343
QY 340 FSROKVSHTTFGHSHLCLGQHLARREIIVTLKWLTRIPDPSIA-PGAIQHKSG-IVS 397
DB 344 WDRPARHHLAFGFGVHQCGLQNLARAEILDIAIRTLFERLPGLRLAVPAHEIRHKPGDTIQ 403
QY 398 GVQALPLVW 406
DB 404 GLLDLPVAV 412

RESULT 13
PCT-US92-10885-11
Sequence 11, Application PC/TUS9210885
GENERAL INFORMATION:
APPLICANT: SARIASLANI, SIMA
TITLE OF INVENTION: CONSTITUTIVE
TITLE OF INVENTION: EXPRESSION OF P450SOY
TITLE OF INVENTION: AND FERRODIXIN-SOY IN
TITLE OF INVENTION: STREPTOMYCES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESSEE: E. I. DU PONT DE NEMOURS
ADDRESSEE: AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: USA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch,
MEDIUM TYPE: 1.0 MB

```

; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh System, 6.0
; SOFTWARE: Microsoft Word, 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10885
; FILING DATE: 19921216
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: GALLEGOS, R. THOMAS
; REGISTRATION NUMBER: 32,692
; REFERENCE/DOCKET NUMBER: CR-9000-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-7342
; TELEFAX: 302-892-7949
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 412 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; PCT-US92-10885-11

```

```

Query Match 14.5%; Score 316; DB 5; Length 412;
Best Local Similarity 27.2%; Pred. No. 3.1e-24;
Matches 84; Conservative 56; Mismatches 163; Indels 6; Gaps 5;

QY 103 MDPBQRQFALANQVGMVVDKLENRIQELACSLIESLRPQG-QCNFTEDYAEPPPIR 161
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
105 VDEPHNTQRMILPTFSVRKIGALRPRIQETVDRLDLMERQGPFAELVSFAFALPVPSM 164
QY 162 IFVLLAGLPEEDIPHLKYLTDQTRPDGSMTEAEAKALDYLIPIIQEQRQKPGTDAIS 221
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
165 VICALLGVFADHAFPERQRLRGSGADVNPARDELEYIGALIDRRAREPGCGLLD 224
QY 222 VANGQVNGRPITSDAKRCGLLVGLDITVNVNFLSFEFLAKSPEHQELIERPERI 281
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
225 ELTHRDHDPDVREQLAVAFVILLIAGHETTANWISLGTFTLLSHPEQLAALRAGTST 284
QY 282 PAACEELLRFSLVADG--RILTSDFYFHGVQLKKGQIQLPQLMGLDERKNCACPMHVD 339
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
285 AVVVEELL-RFLSIAEGLQRLATEDMEVDGATIRKGGVVFSTSLINRDADVFPPRAETLD 343
QY 340 FSRQKVSHTTFHGSHLCLGQHLARREIIVTLKWLTRIPDFSTA-PGAQIQHKSQ-IVS 397
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
344 WDPARHHLAFGFGVHQCGLNARAEIDAMRTLPERLPGLRUAVPAHEERKPGGTIQ 403
QY 398 GVOALPLVW 406
Db : : : : :
404 GLLDLPVW 412

```

```

RESULT 14
US-09-335-409-8
; Sequence 8, Application: US/09335409
; Patent No. 6121329
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335,409
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Sorangium cellulosum

```

```

US-09-335-409-8
Query Match 12.5%; Score 271.5; DB 3; Length 419;
Best Local Similarity 23.8%; Pred. No. 1.4e-19;
Matches 100; Conservative 67; Mismatches 179; Indels 75; Gaps 14;

QY 6 QSNANLAPLPHVPEHLVCEFDWYNPSNLSAGVQE---AWAVLQESNVVDLVWTRCNGS 61
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
3 QEQANQSETKP-----AFCKPFP-----GYAEDPPFAIRLREA-TPIFYWD--EGR 48
QY 62 HWIATRGQ-----LIREAYETVYRHFSECCPPIPREAGEAYDFITSDMPDPQR 109
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
49 SWLTRYHDYSAVFRDERFAVSREMESSAEPSSAIP-----ELSDMKKYGLGFLPPEHDA 104
QY 110 QFRALANQVGMVVDKLENRIQELACSLIESLRPQGQCNFTEDYAEPPPIRIFMLLAGL 169
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
105 KVRKLVNPSFTSRAIDLRLAEIQRITVDQLDARSQGEFFVDVRYDAEGIPMAISALLKV 164
QY 170 PEEDIPHLKYLTDQTRPDGSMTEAEAKALDYLIPI-----I 207
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
165 PAE-----CDEKFRFGSAT-----ARALGVGLVPOVDEETKTLVASVTEGLALLHDV 212
QY 208 IEORRQKP-GTDAISIVANGQVNGRPITSDAKRCGLLVGLDITVNVNFLSFEFLAK 266
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
213 LDERRRNPLENDVLTMLLQAEADGSRSLTKELVALVGAIIAAGTTTIVLIAFAVLNLLR 272
QY 267 SPEHQELIERPERIPACBELIRRFSLVADG--RILTSDFYFHGVQLKKGDOI--LLPQ 322
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
273 SPEALELVKAEPPGLMRNALDEVLRDNILRIGTGVRFARQDLEYCGASIKKGMVELLIPS 332
QY 323 MLSGLDERKNCACPMHVDPSRQKVSHTTFHGSHLCLGQHLARREIIVTLKWLTRIPDFS 382
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
333 ALR--DGVTSRPPVDFVRDRTGSLAYGRPHVCPGVSILARLAEIAVGIFRFPENK 390
QY 383 I 383
Db :
391 L 391

RESULT 15
US-09-413-814-71
; Sequence 71, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bloeker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hefle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
; FILE REFERENCE: PCT/US 99/23535
; CURRENT APPLICATION NUMBER: US/09/413,814
; CURRENT FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; EARLIER FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 71
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-413-814-71

```

```

Query Match 12.5%; Score 271.5; DB 3; Length 419;
Best Local Similarity 23.8%; Pred. No. 1.4e-19;
Matches 100; Conservative 67; Mismatches 179; Indels 75; Gaps 14;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 5, 2004, 18:53:57 ; Search time 34.974 Seconds
(without alignments)
3108.883 Million cell updates/sec

Title: US-09-246-451A-12

Perfect score: 2179

Sequence: 1 TRETQSNANLAPLPVHVE.....IVSGVQALPLVWDPTTKAV 414

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 26263353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 5%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2179	100.0	414	15	US-10-453-104-12
2	2173	99.7	414	15	US-10-453-104-11
3	2169	99.5	414	15	US-10-453-104-2
4	2162	99.2	414	15	US-10-453-104-13
5	383.5	17.6	404	14	US-10-214-446-50
6	382.5	17.6	404	14	US-10-214-446-40
7	378	17.3	399	14	US-10-314-657-37
8	373	17.1	416	14	US-10-156-761-14997
9	371.5	17.0	416	9	US-09-361-289-39
10	371.5	17.0	416	9	US-09-860-846-39
11	371.5	17.0	416	10	US-09-988-384B-39
12	371.5	17.0	416	10	US-09-836-821-39
13	371.5	17.0	416	10	US-09-793-708-18
14	371.5	17.0	416	14	US-10-201-365-13
15	371.5	17.0	416	14	US-10-160-539-18

```

16 371.5 17.0 416 14 US-10-271-889-39
17 369.5 17.0 411 14 US-10-156-761-8376
18 364 16.7 393 14 US-10-156-761-9525
19 363 16.7 405 14 US-10-214-446-38
20 362 16.6 425 14 US-10-214-446-20
21 357.5 16.4 399 14 US-10-156-761-9914
22 357.5 16.4 457 14 US-10-156-761-11073
23 356 16.3 418 12 US-10-389-647-559
24 344.5 15.8 408 14 US-10-214-446-4
25 339 15.6 409 15 US-10-458-201-12
26 338 15.5 392 14 US-10-214-446-32
27 334.5 15.4 388 14 US-10-156-761-13776
28 334.5 15.4 406 14 US-10-214-446-2
29 327 15.0 404 14 US-10-214-446-16
30 324.5 14.9 396 14 US-10-214-446-14
31 324 14.9 428 14 US-10-201-213-6
32 322.5 14.8 404 14 US-10-156-761-14659
33 322 14.8 401 14 US-10-156-761-8710
34 319 14.6 412 14 US-10-214-446-36
35 312 14.3 421 14 US-10-156-761-9703
36 309 14.2 418 14 US-10-214-446-22
37 304.5 14.0 402 14 US-10-205-032-8
38 304.5 14.0 404 14 US-10-156-761-10431
39 304 14.0 475 14 US-10-145-415-22
40 303 13.9 470 14 US-10-145-415-6
41 301.5 13.8 400 14 US-10-314-657-62
42 301.5 13.8 430 9 US-09-738-626-4117
43 301 13.8 399 14 US-10-156-761-7959
44 301 13.8 429 14 US-10-145-415-14
45 300.5 13.8 415 14 US-10-214-446-46

```

ALIGNMENTS

```

RESULT 1
US-10-453-104-12
; Sequence 12, Application US/10453104
; Publication No. US20030207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Hyun Joo
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
; FILE REFERENCE: 4058/1E827-US3
; CURRENT APPLICATION NUMBER: US/10/453.104
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 09/661,093
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 09/246,451
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/086,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/106,834
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Mutant M7-6H
US-10-453-104-12

```

Query Match 100.0%; Score 2179; DB 15; Length 414;
Best Local Similarity 100.0%; Pred. No. 6; Seq-ID 12;
Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTETIQSNANLAPLPHVPEHLVDFDMYNPSNLISAGVQEAVALQESNVPDLVWTRCNG 60
DB 1 TTETIQSNANLAPLPHVPEHLVDFDMYNPSNLISAGVQEAVALQESNVPDLVWTRCNG 60
QY 61 GHWIATRGQILREAYEDYRHSFSECPPIPREAGAYDFIPTSMPPPPQRFALANQVWG 120
DB 61 GHWIATRGQILREAYEDYRHSFSECPPIPREAGAYDFIPTSMPPPPQRFALANQVWG 120
QY 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPPFIRIFMLLAGLPEDIPHLKYL 180
DB 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPPFIRIFMLLAGLPEDIPHLKYL 180
QY 181 TDQWTRPDGSMTPAEAKEALYDYLIPITIEORRQKPGTDAISIVANGVNGRPITSDAKR 240
DB 181 TDQWTRPDGSMTPAEAKEALYDYLIPITIEORRQKPGTDAISIVANGVNGRPITSDAKR 240
QY 241 MCGLLLVGGGLDVTWNFLSFMELAKSPERQELIERPELIPAAACELLRRFSIVADGRI 300
DB 241 MCGLLLVGGGLDVTWNFLSFMELAKSPERQELIERPELIPAAACELLRRFSIVADGRI 300
QY 301 LTSDFEFGVOLKKGQDILLPQMLSGLDERKNACPMHVDFSRQKVSHTTFEGSHLCLGQ 360
DB 301 LTSDFEFGVOLKKGQDILLPQMLSGLDERKNACPMHVDFSRQKVSHTTFEGSHLCLGQ 360
QY 361 HLAARREIIVTLKEWLTRIPDFSAPGAQIQKSGIVSGVQALPLVMDPATTKAV 414
DB 361 HLAARREIIVTLKEWLTRIPDFSAPGAQIQKSGIVSGVQALPLVMDPATTKAV 414

RESULT 2

US-10-453-104-11
; Sequence 11, Application US/10453104
; Publication No. US20030207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Hyun Joo
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
; FILE REFERENCE: 4058/1827-US3
; CURRENT APPLICATION NUMBER: US/10/453,104
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 09/661,093
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 09/246,451
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/086,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/106,834
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant M7-4H
US-10-453-104-11

Query Match 99.7%; Score 2173; DB 15; Length 414;
Best Local Similarity 99.8%; Pred. No. 2.6e-211;
Matches 413; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTETIQSNANLAPLPHVPEHLVDFDMYNPSNLISAGVQEAVALQESNVPDLVWTRCNG 60
DB 1 TTETIQSNANLAPLPHVPEHLVDFDMYNPSNLISAGVQEAVALQESNVPDLVWTRCNG 60
QY 61 GHWIATRGQILREAYEDYRHSFSECPPIPREAGAYDFIPTSMPPPPQRFALANQVWG 120

DB 61 GHWIATRGQILREAYEDYRHSFSECPPIPREAGAYDFIPTSMPPPPQRFALANQVWG 120
QY 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPPFIRIFMLLAGLPEDIPHLKYL 180
DB 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPPFIRIFMLLAGLPEDIPHLKYL 180
QY 181 TDQWTRPDGSMTPAEAKEALYDYLIPITIEORRQKPGTDAISIVANGVNGRPITSDAKR 240
DB 181 TDQWTRPDGSMTPAEAKEALYDYLIPITIEORRQKPGTDAISIVANGVNGRPITSDAKR 240
QY 241 MCGLLLVGGGLDVTWNFLSFMELAKSPERQELIERPELIPAAACELLRRFSIVADGRI 300
DB 241 MCGLLLVGGGLDVTWNFLSFMELAKSPERQELIERPELIPAAACELLRRFSIVADGRI 300
QY 301 LTSDFEFGVOLKKGQDILLPQMLSGLDERKNACPMHVDFSRQKVSHTTFEGSHLCLGQ 360
DB 301 LTSDFEFGVOLKKGQDILLPQMLSGLDERKNACPMHVDFSRQKVSHTTFEGSHLCLGQ 360
QY 361 HLAARREIIVTLKEWLTRIPDFSAPGAQIQKSGIVSGVQALPLVMDPATTKAV 414
DB 361 HLAARREIIVTLKEWLTRIPDFSAPGAQIQKSGIVSGVQALPLVMDPATTKAV 414

RESULT 3

US-10-453-104-2
; Sequence 2, Application US/10453104
; Publication No. US20030207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Hyun Joo
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
; FILE REFERENCE: 4058/1827-US3
; CURRENT APPLICATION NUMBER: US/10/453,104
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 09/661,093
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 09/246,451
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/086,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/106,834
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 414
; TYPE: PRT
; ORGANISM: P. Putida
US-10-453-104-2

Query Match 99.5%; Score 2169; DB 15; Length 414;
Best Local Similarity 99.5%; Pred. No. 6.7e-211;
Matches 412; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTETIQSNANLAPLPHVPEHLVDFDMYNPSNLISAGVQEAVALQESNVPDLVWTRCNG 60
DB 1 TTETIQSNANLAPLPHVPEHLVDFDMYNPSNLISAGVQEAVALQESNVPDLVWTRCNG 60
QY 61 GHWIATRGQILREAYEDYRHSFSECPPIPREAGAYDFIPTSMPPPPQRFALANQVWG 120
DB 61 GHWIATRGQILREAYEDYRHSFSECPPIPREAGAYDFIPTSMPPPPQRFALANQVWG 120
QY 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPPFIRIFMLLAGLPEDIPHLKYL 180
DB 121 MPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPPFIRIFMLLAGLPEDIPHLKYL 180
QY 181 TDQWTRPDGSMTPAEAKEALYDYLIPITIEORRQKPGTDAISIVANGVNGRPITSDAKR 240

Db 181 TDQMPDGSMTFAEAKAALDYLLPIIEQRKPGCTDAISIVANGQVNGRPITSDEAKR 240
QY 241 MCGLLLVGGGLDVTNNFLSFSMEFLAKSPHRRQELIERPELIPACBELLRRFSLVADGRI 300
Db 241 MCGLLLVGGGLDVTNNFLSFSMEFLAKSPHRRQELIERPERIPACBELLRRFSLVADGRI 300
QY 301 LTSDYEFHGVQVKKGQDQILLPQMLSGLDERENACPMHVDFSRQKVSHTTTFGHGSHLCIGQ 360
Db 301 LTSDYEFHGVQVKKGQDQILLPQMLSGLDERENACPMHVDFSRQKVSHTTTFGHGSHLCIGQ 360
QY 361 HLAARREIIVTLKEWLTRIPDPSIAPGAQIQHKSIGVSGVQALPLVMDPATTKAV 414
Db 361 HLAARREIIVTLKEWLTRIPDPSIAPGAQIQHKSIGVSGVQALPLVMDPATTKAV 414
RESULT 4
US-10-453-104-13
; Sequence 13, Application US/10453104
; Publication No. US20030207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Eyun Joo
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
; FILE REFERENCE: 4C58/E827-US3
; CURRENT APPLICATION NUMBER: US/10/453,104
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 09/661,093
; PRIOR FILING DATE: 2000-09-23
; PRIOR APPLICATION NUMBER: US 09/246,451
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/086,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/106,834
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant M7-8H
US-10-453-104-13

Query Match 99.2%; Score 2162; DB 15; Length 414;
Best Local Similarity 99.5%; Pred. No. 3.5e-210;
Matches 412; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TTETIQSNANLAPLPHVPEHLVDFEDMNPNSLSAGVQAEAWVLCSNVPDLVWTRCNG 60
Db 1 TTETIQSNANLAPLPHVPEHLVDFEDMNPNSLSAGVQAEAWVLCSNVPDLVWTRCNG 60
QY 61 GHWIATRGQILREAYEDYRHFSSECFPIPREAGEAYDPIPTSMCPQORFALANQVVG 120
Db 61 GHWIATRGQILREAYEDYRHFSSECFPIPREAGEAYDPIPTSMCPQORFALANQVVG 120
QY 121 MPVVDKLENRIQELACSLIESLRPOGCTEYARPPRIEMLLAGLPEEDIPLHKYL 180
Db 121 MPVVDKLENRIQELACSLIESLRPOGCTEYARPPRIEMLLAGLPEEDIPLHKYL 180
QY 181 TDQMPDGSMTFAEAKAALDYLLPIIEQRKPGCTDAISIVANGQVNGRPITSDEAKR 240
Db 181 TDQMPDGSMTFAEAKAALDYLLPIIEQRKPGCTDAISIVANGQVNGRPITSDEAKR 240
QY 241 MCGLLLVGGGLDVTNNFLSFSMEFLAKSPHRRQELIERPELIPACBELLRRFSLVADGRI 300
Db 241 MCGLLLVGGGLDVTNNFLSFSMEFLAKSPHRRQELIERPERIPACBELLRRFSLVADGRI 300

QY 301 LTSDYEFHGVQVKKGQDQILLPQMLSGLDERENACPMHVDFSRQKVSHTTTFGHGSHLCIGQ 360
Db 301 LTSDYEFHGVQVKKGQDQILLPQMLSGLDERENACPMHVDFSRQKVSHTTTFGHGSHLCIGQ 360
QY 361 HLAARREIIVTLKEWLTRIPDPSIAPGAQIQHKSIGVSGVQALPLVMDPATTKAV 414
Db 361 HLAARREIIVTLKEWLTRIPDPSIAPGAQIQHKSIGVSGVQALPLVMDPATTKAV 414
RESULT 5
US-10-214-446-50
; Sequence 50, Application US/10214446
; Publication No. US20030180742A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, David
; APPLICANT: Burk, Mark J.
; APPLICANT: Hitchman, Tim
; APPLICANT: Pujol, Catherine
; APPLICANT: Richardson, Toby
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: F450 ENZYMES, NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: THEM AND METHODS OF MAKING AND USING THEM
; FILE REFERENCE: 09010-500001
; CURRENT APPLICATION NUMBER: US/10/214,446
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 60/309,497
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Bacterial
US-10-214-446-50
Query Match 17.6%; Score 383.5; DB 14; Length 404;
Best Local Similarity 27.9%; Pred. No. 8.4e-30;
Matches 98; Conservative 64; Mismatches 156; Indels 33; Gaps 5;

QY 60 GHWIATRGQILREAYEDYRHFSSECFPIPREAGEAYDPIPTSM----- 103
Db 32 GRTWFLPHHADIRTLALRDERFSAS-----RKAGGFVNQPPAEVRPEFAFNEAISRWIV 85
QY 104 --DPPEQORFALANOVGVVVKLENRIQELACSLIESLRPOGCTEYARPPRI 161
Db 86 LHDQPEHRQRLQMLQQGFTRLITTEPKIORVCDLIDAFVKKGSTEFMTETIAHPFAK 145
QY 162 IFMLLAGLPEEDIPLHKYLTDQMPDGSMTFAEAKAALDYLLPIIEQRK 213
Db 146 VIAEMLVNPEDYPAFVWSEDLINPAGSLRPTLEMFRAAQDGLAMMDYFALLPERRE 205
QY 214 KPQTDALISIVANGQVNGRPITSDEAKMCGLLLVGGGLDVTNNFLSFSMEFLAKSPHROE 273
Db 206 NPGDGLVSLLSAESEGEWMTAEQVLANCTQIIVAGHETTRNLVANGVELLRYPBQSA 265
QY 274 LISEPELIPACBELLRRFS-LVADGRILLSDYEFHGVQVKKGQDQILLPQMLSGLDEREN 332
Db 266 IESRPELMPASVREIMHMFESPLQIRKVAAREDEFFGAEVREGGLVMLGSLNARDPEAF 325
QY 333 ACMHVDFSRQKVSHTTTFGHGSHLCIGQLARREIIVTLKEWLTRIPDPSI 383
Db 326 DDPDPTDLTENPTGHLAAGVGPVCVGAALAELEGQVSFRTLLDLPLGJEL 376

RESULT 6
US-10-214-446-50
; Sequence 40, Application US/10214446
; Publication No. US20030180742A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, David
; APPLICANT: Burk, Mark J.
; APPLICANT: Hitchman, Tim
; APPLICANT: Pujol, Catherine

```

; APPLICANT: Richardson, Toby
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: P450 ENZYMES, NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 09010-500001
; CURRENT APPLICATION NUMBER: US/10/214.446
; PRIOR FILING DATE: 2002-08-05
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Bacterial
; ORGANISM: Streptomyces atroolivaceus
US-10-214-446-40

Query Match 17.6%; Score 382.5; DB 14; Length 404;
Best Local Similarity 30.3%; Pred. No. 1.1e-29;
Matches 120; Conservative 58; Mismatches 187; Indels 31; Gaps 15;

QY 31 PSNLSAGVQAMAVLQESNVDPDLVWTRCNGHW--WIATRGQLIREAYEDYRHFSSECPFI 88
Db 20 PFNEADGISLADAYEEAREQFGLLRVMAYCEPAWLATR-----YADARLVLGDRRFS 72
QY 89 -----PREAGEAYDFTPTSDPPEQRFALANQVGVMPVVDKLENRIQELACSLI 139
Db 73 RAEGARHDEPRQSGRDSGLSDMDPDHTRLTFLVAKFTMHQVKEKLRPAVRLEADELI 132
QY 140 ESLRPGQ-QCNFTEDYAEPPPIRIFMLLAGLPEEDI PHLKYLTDQMTDPGSMT---FAE 195
Db 133 DKVATGAPVDLVSEEPALPVGVVICQLLVGVVEEERFRFAWSD-AALSTSLTAEFFDA 191
QY 196 AKALVDYLPITIQRQKPGTDAISIVANGQVNGRZITSDAKRMCGLLVGGLOTVWN 255
Db 192 NQBELRAYMGLIEDHRAHPREDITGLIEARDDDRLTEQELVDLCVGLVAGHETAT 251
QY 256 FLSSFEFLAKSPHROELIERPELIPAAABELLRSELSVADG---RLTSDYEPHGVQ 311
Db 252 QIFNVVTLDRPEQWNRLEDPELVTAVEELM-RFVPLGSGASFPRYATEDVEVGGTL 310
QY 312 LKXGQDILLPQMLSG-LDERKNACPMHVDFFGRQKVSHTTFGGHSHLCGLQHLARREIIVTL 371
Db 311 VRAGEPVLVAGAAANRDPAPDPAPCELDLAREGNQHLGFGHGVHHCILGAPLARLEQAL 370
QY 372 KENWTRIPDSIAPGAQIQHKS-GIVSGVQALPLVW 406
Db 371 GALLRRLPGLRIA--GDIEWKTQMLVRGPRTLPLVW 404

RESULT 7
US-10-314-657-37
; Sequence 37, Application US/10314657
; Publication No. US20030175888A1
; GENERAL INFORMATION:
; APPLICANT: SHEN, Ben
; APPLICANT: CHENG, Yi-Qiang
; APPLICANT: TANG, Gong-Li
; TITLE OF INVENTION: Discrete Acyltransferases Associated with Type I Polyketide
; FILE REFERENCE: 054030-3021
; CURRENT APPLICATION NUMBER: US/10/314.657
; CURRENT FILING DATE: 2002-12-09
; PRIOR FILING DATE: 2002-03-22
; PRIOR FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIt version 3.2
; SEQ ID NO 37
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Streptomyces atroolivaceus

US-10-314-657-37
Query Match 17.3%; Score 378; DB 14; Length 399;
Best Local Similarity 28.3%; Pred. No. 3e-29;
Matches 111; Conservative 70; Mismatches 175; Indels 36; Gaps 9;

QY 38 VQEAVALQESNVDPDLVWTRCNGHW-ATRGQLIREAYEDYRHFSSECPFI 94
Db 21 IHKFAELRETDPLARVLPYGGEGKWYTR-----YDDVRAANSDFRFSQAQIGETP 73
QY 95 -----AYDFTPTSDPPEQRFALANQVGVMPVVDKLENRIQELACSLIESLRPGQ 147
Db 74 RTTPLARRSDTI-LSLDPPEHTRRLRLSKAFTARRMGAMQSWLEELFAGLLDGVERTGH 132
QY 148 -CNFTEDYAEPPPIRIFMLLAGLPEEDI PHLKYLTDQMTDPGSMTAEAKE-----A 199
Db 133 PADIVRLAOFPTTAVICRLGVVEYEDRGRFQHWSEVI-----MSTTAYSKEEAVSADAS 187
QY 200 LYDLILPIIEQRQKPGTDAISIVANGQVNGRPTTSDAKRMCGLLVGGLOTVWNFLSF 259
Db 188 IRVLADIVSARRAAPHDDLGLVLSARDDDRLTEDELIITFGVTLVAGHETSAHQNGN 247
QY 260 SMEFLAKSPHROELIERPELIPAAABELLRFSL---VADGRILTSYEPHGVOLKXGD 316
Db 248 MVYALLTHEQDLSLRQPELLPRAVELLAFVPLGVNGVGNARIALEDELVELSGGTVRAGE 307
QY 317 QILLPQMLSG-LDERKNACPMHVDFFSRQKVSHTTFGGHSHLCGLQHLARREIIVTLKEWLT 376
Db 308 GVVAANVNAARDPRAPDDPDRDLITREKNPHLAFGHGAHYCLGAGLAQMLARVLAIGGLE 367
QY 377 RIPDPSIA-PGAQIQHKS-GIVSGVQALPLVW 406
Db 368 RFPGLRLAVPADQVEMWKTGGLFRGQRLPIAW 399

RESULT 8
US-10-156-761-14997
; Sequence 14997, Application US/10156761
; Publication No. US20030119318A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156.761
; CURRENT FILING DATE: 2002-05-29
; PRIOR FILING DATE: 2001-05-30
; PRIOR FILING DATE: 2001-05-30
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15129
; SEQ ID NO 14997
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14997

Query Match 17.1%; Score 373; DB 14; Length 416;
Best Local Similarity 29.4%; Pred. No. 1e-28;
Matches 105; Conservative 56; Mismatches 180; Indels 16; Gaps 8;

QY 63 WIATRGQLIREAYEDYRHFSSECPFIER--EAGEAYDFTPT-----SMDPPEQRFAL 114
Db 63 WYVTGHAARALLSDQLSSDRT--LRFPPATFERFAVTRRVRVALLGVDDPDEHRTQRM 120
QY 115 ANQVGVMPVVDKLENRIQELACSLIESLRPGQ-OCNFTEDYAEPPPIRIFMLLAGLPEED 173
Db 121 LVPSFTLKRAAALRFRIQETVVDGLDAMEAQGPPELVSAFALFPLPSMVICALLGVPPYAD 180
```


QY 174 IPHLXYLTQMTDPGSMFAEAKALYDYLPIIQRQKPGTDAISIVANGVNGRPI 233
 Db 181 HFFESQSRRLRGPIAEVQDRAQDDLYALIDRKREPGDGLDDLIQQLARGTV 240
 QY 234 TDEAKRMCGLLAVGSLCTVNFSLFSMEFLAKSPHROBELIRPELIPAAECLELRRFS 293
 Db 241 DRAELVSLATLLIAGHEITANNISLGTFTLLRHPEQLAEIRAEGLMPAAVEELL-REL 299
 QY 294 LVADG--RLTSDYEHGVLQKGDQILLPQMLSGLDERKNACPMHVDPSRQKVSHTTFF 351
 Db 300 STADGHLRVATEDIEVAGTTIRADEGVFATSINRDAAGFAEPDADLMHRSARHHVAFG 359
 QY 352 HGHSLCLOGLARREIIVTKELARIPDSI-APGAQIOHKSG-IVSGVQALPLVW 406
 Db 360 FGIHQCLQNLARAEAMEIALGT-FEHLPGRLAARPADEIPFKGDTIQGMLBPTVW 416

RESULT 9

US-09-861-289-39
 ; Sequence 39, Application US/09861289
 ; Patent No. US20020112897A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sherman, D.H.
 ; APPLICANT: Liu, H.
 ; APPLICANT: Xue, Y.
 ; APPLICANT: Zhao, L.
 ; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
 ; FILE REFERENCE: 600.438US1
 ; CURRENT APPLICATION NUMBER: US/09/861,289
 ; PRIOR FILING DATE: 2001-05-18
 ; PRIOR APPLICATION NUMBER: 09/105,537
 ; PRIOR FILING DATE: 1998-06-26
 ; NUMBER OF SEQ ID NOS: 43
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 39
 ; LENGTH: 416
 ; TYPE: PRT
 ; ORGANISM: Streptomyces venezuelae
 US-09-861-289-39

Query Match 17.0%; Score 371.5; DB 9; Length 416;
 Best Local Similarity 28.6%; Pred. No. 1.4e-28;
 Matches 98; Conservative 62; Mismatches 162; Indels 21; Gaps 7;
 QY 76 EDYRHSSECPFTPREAGEAYDPIPTSMDDPEQRFALANQVGVMPVVDKLENRIOELA 135
 Db 72 KDWK--NSTTPTTEAEALNHNMLE--DPRHTRLRKLVARERTMRVVELLRPRVQEI 127
 QY 136 CSLIESL--RPOGQCNFTDYAPFPPIRIFMLLAGLPEEDIPHLKYLTDQMTDPGSMTF 193
 Db 128 DGLVDAMLAAPDGRADIMESLAWPLPITVISELLGVPEPDRAFRVWTDAFVFPDPAQA 187
 QY 194 ABAKEALYDYLPIIQRQKPGTDAIS-IVANGVNGRPIITSDAKRMCGLLVGLGLOT 252
 Db 188 QTMAEMSGYLSRLIDSKRGQDGEDLLSA-LVRSDEGSLTSEELGMAHILLVAGHET 247
 QY 253 VVNFSLFSMEFLAKSPHROBELIRPELIPAAECLELLRRLSLVADGRILTSDFEYF---- 307
 Db 248 TVNLIANGMYALLSHPDQALRADMTLLDGAVEEMLR-----YEGVESATYFFPVEPV 302
 QY 308 --HGVOLKKGDDQILLPQMLSGLDERKNACPMHVDPSRQKVSHTTFFGHSHLCLOHLARR 365
 Db 303 DLDGTVIPAGDTVLVWLADAHRTPERPDPHRRDIRDRTAGHLAAGHGHIFCIGAPLARL 362
 QY 366 EIIVTLKWLTRIPDS--IAPGAQIOHKSGIVSGVQALPLVW 406
 Db 363 EARIAVRALLERCPDIALDVSPGELVWYPNMIRGLKALPIRW 405

RESULT 10

US-09-860-846-39
 ; Sequence 39, Application US/09860846

; Patent No. US20020164742A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sherman, D.H.
 ; APPLICANT: Liu, H.
 ; APPLICANT: Xue, Y.
 ; APPLICANT: Zhao, L.
 ; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
 ; FILE REFERENCE: 600.438US1
 ; CURRENT APPLICATION NUMBER: US/09/860,846
 ; CURRENT FILING DATE: 2001-05-18
 ; PRIOR APPLICATION NUMBER: 09/105,537
 ; PRIOR FILING DATE: 1998-06-26
 ; NUMBER OF SEQ ID NOS: 43
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 39
 ; LENGTH: 416
 ; TYPE: PRT
 ; ORGANISM: Streptomyces venezuelae
 US-09-860-846-39

Query Match 17.0%; Score 371.5; DB 9; Length 416;
 Best Local Similarity 28.6%; Pred. No. 1.4e-28;
 Matches 98; Conservative 62; Mismatches 162; Indels 21; Gaps 7;
 QY 76 EDYRHSSECPFTPREAGEAYDPIPTSMDDPEQRFALANQVGVMPVVDKLENRIOELA 135
 Db 72 KDWK--NSTTPTTEAEALNHNMLE--DPRHTRLRKLVARERTMRVVELLRPRVQEI 127
 QY 136 CSLIESL--RPOGQCNFTDYAPFPPIRIFMLLAGLPEEDIPHLKYLTDQMTDPGSMTF 193
 Db 128 DGLVDAMLAAPDGRADIMESLAWPLPITVISELLGVPEPDRAFRVWTDAFVFPDPAQA 187
 QY 194 ABAKEALYDYLPIIQRQKPGTDAIS-IVANGVNGRPIITSDAKRMCGLLVGLGLOT 252
 Db 188 QTMAEMSGYLSRLIDSKRGQDGEDLLSA-LVRSDEGSLTSEELGMAHILLVAGHET 247
 QY 253 VVNFSLFSMEFLAKSPHROBELIRPELIPAAECLELLRRLSLVADGRILTSDFEYF---- 307
 Db 248 TVNLIANGMYALLSHPDQALRADMTLLDGAVEEMLR-----YEGVESATYFFPVEPV 302
 QY 308 --HGVOLKKGDDQILLPQMLSGLDERKNACPMHVDPSRQKVSHTTFFGHSHLCLOHLARR 365
 Db 303 DLDGTVIPAGDTVLVWLADAHRTPERPDPHRRDIRDRTAGHLAAGHGHIFCIGAPLARL 362
 QY 366 EIIVTLKWLTRIPDS--IAPGAQIOHKSGIVSGVQALPLVW 406
 Db 363 EARIAVRALLERCPDIALDVSPGELVWYPNMIRGLKALPIRW 405

RESULT 11

US-09-988-384B-39
 ; Sequence 39, Application US/0988384B
 ; Publication No. US20030073824A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sherman, D.H.
 ; APPLICANT: Liu, H.
 ; APPLICANT: Xue, Y.
 ; APPLICANT: Zhao, L.
 ; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
 ; FILE REFERENCE: 600.536US1
 ; CURRENT APPLICATION NUMBER: US/09/988,384B
 ; CURRENT FILING DATE: 2001-11-19
 ; PRIOR APPLICATION NUMBER: PCT/US99/14398
 ; PRIOR FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: US 09/105,537
 ; PRIOR FILING DATE: 1998-06-26
 ; NUMBER OF SEQ ID NOS: 53
 ; SEQ ID NO 39
 ; LENGTH: 4-6
 ; TYPE: PRT
 ; ORGANISM: Streptomyces venezuelae
 US-09-988-384B-39

; Sequence 13, Application US/10201365

; Publication No. US20030148463A1

; GENERAL INFORMATION: Ashley, Gary

; APPLICANT: ASHLEY, Gary

; APPLICANT: BETLACH, Melanie C.

; APPLICANT: BETLACH, Mary

; APPLICANT: MCDANIEL, Robert

; APPLICANT: TANG, Li

; TITLE OF INVENTION: COMBINATORIAL POLYKETIDE LIBRARIES PRODUCED USING A MODULAR

; FILE REFERENCE: 300622002103

; CURRENT APPLICATION NUMBER: US/10/201,365

; CURRENT FILING DATE: 2002-07-22

; PRIOR APPLICATION NUMBER: US 09/141,908

; PRIOR FILING DATE: 1998-08-28

; PRIOR APPLICATION NUMBER: US 09/073,538

; PRIOR FILING DATE: 1998-05-06

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 13

; LENGTH: 416

; TYPE: PRT

; ORGANISM: Streptomyces venezuelae

; US-10-201-365-13

Query Match 17.0%; Score 371.5; DB 14; Length 416;

Best Local Similarity 28.6%; Pred. No. 1.4e-28;

Matches 98; Conservative 62; Mismatches 162; Indels 21; Gaps 7;

Qy 76 EDYRHSSECFPIPREAGEAYDFTSMDPPEQRFALANQVGMVVDKLENRIQELA 135

Db 72 KQWR--NSTTPTAEAAALNNMLES--DPRHTRLRKLVAREFTMRVVELLRPRVQELV 127

Qy 136 CSLIESL--RPOGQCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYLTDQWTRPDGSMTF 193

Db 128 DGLVDAMLAAPDGRADLMESLAWPLPTIV--SELLGVPEPRAAFRVWTDAPVFPDDPAQA 187

Qy 194 AEAKEALYDYLPIIPIEORRQKPGTDAIS--IVANGQVNGRPITSDAKRMCGLLLVGGIDT 252

Db 188 QTAMAEYSGVLSRLIDSKRGQDGEDLSALVTSDEGSRITSEELGMAHILLVAGHET 247

Qy 253 VVNFLSFMEFLAKSPHQRQELIERPELIPAAACELLRRFSLVADGRILTSDFEY----- 307

Db 248 TVNLIANGMYALLSHPPDLAALRADMTLLDGAVEMLR-----YEGPVSATYRFPVEFV 302

Qy 308 --HGVQLKKGQDQILLPQMLSGLDERKNACPMHVDPSQKVSHHTTFHGSHLCLGQHILARR 365

Db 303 DLDTGTVIPAGDTVLVWLADAHRTPERFDPHRTDIRDTAGHLAFGHHGIFCIGAPLARL 362

Qy 366 EIIIVTLKEWLTRIPDFS--IAPGAQIOHKSQIVSGVQALPLVW 406

Db 363 EARIATVALLERCPLDALDVSFGELVWYFNPIMRGLKALPIRW 405

RESULT 15

US-10-160-539-18

; Sequence 18, Application US/10160539

; Publication No. US20030162262A1

; GENERAL INFORMATION: Ashley, Gary

; APPLICANT: ASHLEY, Gary

; APPLICANT: BETLACH, Melanie C.

; APPLICANT: BETLACH, Mary C.

; APPLICANT: MCDANIEL, Robert

; APPLICANT: TANG, Li

; TITLE OF INVENTION: RECOMBINANT NARONOLIDE POLYKETIDE SYNTHASE

; FILE REFERENCE: 300622002120

; CURRENT APPLICATION NUMBER: US/10/160,539

; CURRENT FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: US 09/657,440

; PRIOR FILING DATE: 2000-09-07

; PRIOR APPLICATION NUMBER: 09/320,878

; PRIOR FILING DATE: 1999-05-27

; PRIOR APPLICATION NUMBER: CIP OF 09/141,908

; PRIOR FILING DATE: 1998-08-28

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 18

; LENGTH: 416

; TYPE: PRT

; ORGANISM: Streptomyces venezuelae

; US-10-160-539-18

Query Match 17.0%; Score 371.5; DB 14; Length 416;

Best Local Similarity 28.6%; Pred. No. 1.4e-28;

Matches 98; Conservative 62; Mismatches 162; Indels 21; Gaps 7;

Qy 76 EDYRHSSECFPIPREAGEAYDFTSMDPPEQRFALANQVGMVVDKLENRIQELA 135

Db 72 KQWR--NSTTPTAEAAALNNMLES--DPRHTRLRKLVAREFTMRVVELLRPRVQELV 127

Qy 136 CSLIESL--RPOGQCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYLTDQWTRPDGSMTF 193

Db 128 DGLVDAMLAAPDGRADLMESLAWPLPTIV--SELLGVPEPRAAFRVWTDAPVFPDDPAQA 187

Qy 194 AEAKEALYDYLPIIPIEORRQKPGTDAIS--IVANGQVNGRPITSDAKRMCGLLLVGGIDT 252

Db 188 QTAMAEYSGVLSRLIDSKRGQDGEDLSALVTSDEGSRITSEELGMAHILLVAGHET 247

Qy 253 VVNFLSFMEFLAKSPHQRQELIERPELIPAAACELLRRFSLVADGRILTSDFEY----- 307

Db 248 TVNLIANGMYALLSHPPDLAALRADMTLLDGAVEMLR-----YEGPVSATYRFPVEFV 302

Qy 308 --HGVQLKKGQDQILLPQMLSGLDERKNACPMHVDPSQKVSHHTTFHGSHLCLGQHILARR 365

Db 303 DLDTGTVIPAGDTVLVWLADAHRTPERFDPHRTDIRDTAGHLAFGHHGIFCIGAPLARL 362

Qy 366 EIIIVTLKEWLTRIPDFS--IAPGAQIOHKSQIVSGVQALPLVW 406

Db 363 EARIATVALLERCPLDALDVSFGELVWYFNPIMRGLKALPIRW 405

Search completed: April 6, 2004, 19:14:30

Job time : 34.974 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 13:49:16 ; Search time 14.9588 Seconds
(without alignments)
1428.803 Million cell updates/sec

Title: US-09-346-451A-12

Perfect score: 2179

Sequence: 1 TTTCTCSNANLAPLPHVPE.....IVSGVQALPLYWDP2ATTKAV 414

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51E25971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	371.5	17.0	416	3	US-09-320-878-18
2	371.5	17.0	416	3	US-09-105-537-39
3	371.5	17.0	416	4	US-09-141-908-13
4	371.5	17.0	416	4	US-09-657-440-18
5	355	16.3	437	4	US-09-352-991A-17836
6	346	15.9	403	6	5212296-9
7	339	15.6	409	3	US-09-385-028-12
8	339	15.6	409	4	US-09-726-614-12
9	339	15.6	439	4	US-09-385-040-12
10	336	15.4	436	6	5212296-6
11	323.5	14.8	399	4	US-08-765-907A-10
12	315	14.5	412	1	US-08-102-863-11
13	315	14.5	412	5	PCT-US92-10885-11
14	277.5	12.7	419	3	US-09-335-409-8
15	277.5	12.7	419	3	US-09-413-814-71
16	277.5	12.7	419	4	US-09-568-102-8
17	277.5	12.7	419	4	US-09-567-969-8
18	277.5	12.7	419	4	US-09-568-480-8
19	277.5	12.7	419	4	US-09-568-486-8
20	277.5	12.7	419	4	US-09-568-472-8
21	277.5	12.7	419	4	US-09-567-899-8
22	265	12.2	395	4	US-09-266-965-129
23	240	11.0	468	4	US-09-252-991A-32437
24	174.5	7.3	189	4	US-09-679-279-20
25	160	7.0	422	2	US-09-096-982-5
26	160	7.3	422	2	US-08-653-650A-5
27	160	7.3	474	2	US-09-096-982-8

28	160	7.3	474	2	US-08-653-650A-8	Sequence 8, Appli
29	157	7.2	443	2	US-09-096-982-9	Sequence 9, Appli
30	157	7.2	443	2	US-08-653-650A-9	Sequence 9, Appli
31	155	7.1	422	1	US-08-396-218-2	Sequence 2, Appli
32	155	7.1	422	1	US-08-760-116-2	Sequence 2, Appli
33	135	6.2	512	2	US-08-194-981E-5	Sequence 5, Appli
34	129.5	5.9	382	3	US-09-320-878-7	Sequence 7, Appli
35	129.5	5.9	382	4	US-09-141-908-7	Sequence 7, Appli
36	129.5	5.9	382	4	US-09-657-440-7	Sequence 7, Appli
37	129.5	5.9	402	3	US-09-105-537-22	Sequence 22, Appli
38	129.5	5.9	3782	3	US-09-105-537-4	Sequence 4, Appli
39	127.5	5.9	524	4	US-09-126-420A-24	Sequence 24, Appli
40	127	5.8	516	4	US-09-215-694-16	Sequence 16, Appli
41	120.5	5.5	503	4	US-09-583-447A-2	Sequence 2, Appli
42	120	5.5	503	4	US-09-144-367-2	Sequence 2, Appli
43	120	5.5	504	4	US-09-583-447A-4	Sequence 4, Appli
44	119.5	5.5	504	4	US-09-976-594-642	Sequence 642, App
45	119	5.5	501	3	US-08-906-791-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-320-878-18
; Sequence 18, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: McDaniel, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062202120
; CURRENT FILING DATE: 1999-05-27
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,536
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119,139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100,880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 18
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-320-878-18

Query Match	17.0%	Score 371.5;	DB 3;	Length 416;
Best Local Similarity	28.6%	Pred. No. 7.7e-31;		
Matches	98;	Conservative	62;	Mismatches 162;
Indels	21;	Gaps	7;	
Qy	76	EDYRHSSECPETPREAGHAYDFIPTSDPQOFALANOVVMPVVDKLENRIQELA	135	
Db	72	KDWR--NSTYPLTEAAGALNNHML--DPRHTRUKLVAREFTMRVRLPRVQEV	127	
Qy	136	CSLIESL--RPQCCNFTEYAPFFIRIFMLAGLPEEDIPHLKYLTDQWTRPDGSMTF	193	
Db	128	DGLVDAMLAAPQGRADLMESLAMPITVISELLGVPEPORAARVWTDARVDEDDAQA	187	
Qy	194	AEKAEALYKLLPIIQRQKQKPTDAIS--IVANGQVNGRPITSDEAKRMGLLIVGLDT	252	
Db	188	QTAWAEMSGYLSRLIDSKRGQDGLLSALVRTSDEGSLTSELIGMAHILLVAGHET	247	

```

QY 253 VNFSLFSMEFLAKSPHEHQEIERPELIPAAACEELLRRFSLVADGRILTSDYEF----- 307
Db 248 TVNLIAKXVALLSHPDQALRADMTLLCGAVEMLR-----YEGPVESATYRFPVPV 302
QY 308 --HGVLKXGDOILLPQMLSGLDERKNACPMHVDFSRQKVSHTTFHGSHLCLGQHLARR 365
Db 303 DLGGTVIPAGDVTLVVLADAHRTPEFPDPHRRDRTAGHLAFGHGHCIGAPLARRL 362
QY 366 EIVTLKEMWLRIPDFS--IAPGAQIOHKGSGVGVQALPLVW 406
Db 363 EARIARALLERCEDLALDVSPGELVWVWPNMIRGLKALPIRW 405

RESULT 2
US-09-105-537-33
; Sequence 39, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438USI
; CURRENT APPLICATION NUMBER: US/09/105,537A
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRN
; ORGANISM: Streptomyces venezuelae
US-09-105-537-39

```

```

Query Match 17.0%; Score 371.5; DB 3; Length 416;
Best Local Similarity 28.6%; Pred. No. 7.7e-31;
Matches 98; Conservative 62; Mismatches 162; Indels 21; Gaps 7;

QY 76 EDYRHSFSECFPIPREAGEAYDFPTSDPPEQORFRALANOVGVVVDKLENRIQELA 135
Db 72 KDMR--NSTTPTLTAEEAALNNHML--DPPHTRLRKLVAAREFTMRVE--LRPRVQEV 127
QY 136 CSLIESL--RPGQCNTEDYAEPTPIRIFMLAGLPEDIPHLKYLTDQWTRPDGSMTF 193
Db 128 DGLVDMLAAPDGRADLMESLAWLPITVISELLGVPEPDRAAFVWTDADFVFPDDPAQA 187
QY 194 AEAKALYDYLPIIEORROKPGTDAIS--IVANGQVNGRPITSDAKRMCGLLLVGGDLT 252
Db 188 QTAMAEMSGYLSRLIDSKRGQDGEDLLSALVRTSDGSLTSELLGMHILLVAGHET 247
QY 253 VNFSLFSMEFLAKSPHEHQEIERPELIPAAACEELLRRFSLVADGRILTSDYEF----- 307
Db 248 TVNLIAKXVALLSHPDQALRADMTLLCGAVEMLR-----YEGPVESATYRFPVPV 302
QY 308 --HGVLKXGDOILLPQMLSGLDERKNACPMHVDFSRQKVSHTTFHGSHLCLGQHLARR 365
Db 303 DLGGTVIPAGDVTLVVLADAHRTPEFPDPHRRDRTAGHLAFGHGHCIGAPLARRL 362
QY 366 EIVTLKEMWLRIPDFS--IAPGAQIOHKGSGVGVQALPLVW 406
Db 363 EARIARALLERCEDLALDVSPGELVWVWPNMIRGLKALPIRW 405

```

```

RESULT 3
US-09-141-908-13
; Sequence 13, Application US/09141908
; Patent No. 6503741
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li

```

```

; TITLE OF INVENTION: Combinatorial Polyketide Libraries Produced Using a
; FILE REFERENCE: Modular PKS Gene Cluster as Scaffold
; CURRENT APPLICATION NUMBER: US/09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER FILING DATE: 1998-05-06
; EARLIER FILING DATE: 1998-04-30
; EARLIER FILING DATE: 1998-03-05
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 13
; LENGTH: 416
; TYPE: PRN
; ORGANISM: Streptomyces venezuelae
US-09-141-908-13

Query Match 17.0%; Score 371.5; DB 4; Length 416;
Best Local Similarity 28.6%; Pred. No. 7.7e-31;
Matches 98; Conservative 62; Mismatches 162; Indels 21; Gaps 7;

QY 76 EDYRHSFSECFPIPREAGEAYDFPTSDPPEQORFRALANOVGVVVDKLENRIQELA 135
Db 72 KDMR--NSTTPTLTAEEAALNNHML--DPPHTRLRKLVAAREFTMRVE--LRPRVQEV 127
QY 136 CSLIESL--RPGQCNTEDYAEPTPIRIFMLAGLPEDIPHLKYLTDQWTRPDGSMTF 193
Db 128 DGLVDMLAAPDGRADLMESLAWLPITVISELLGVPEPDRAAFVWTDADFVFPDDPAQA 187
QY 194 AEAKALYDYLPIIEORROKPGTDAIS--IVANGQVNGRPITSDAKRMCGLLLVGGDLT 252
Db 188 QTAMAEMSGYLSRLIDSKRGQDGEDLLSALVRTSDGSLTSELLGMHILLVAGHET 247
QY 253 VNFSLFSMEFLAKSPHEHQEIERPELIPAAACEELLRRFSLVADGRILTSDYEF----- 307
Db 248 TVNLIAKXVALLSHPDQALRADMTLLCGAVEMLR-----YEGPVESATYRFPVPV 302
QY 308 --HGVLKXGDOILLPQMLSGLDERKNACPMHVDFSRQKVSHTTFHGSHLCLGQHLARR 365
Db 303 DLGGTVIPAGDVTLVVLADAHRTPEFPDPHRRDRTAGHLAFGHGHCIGAPLARRL 362
QY 366 EIVTLKEMWLRIPDFS--IAPGAQIOHKGSGVGVQALPLVW 406
Db 363 EARIARALLERCEDLALDVSPGELVWVWPNMIRGLKALPIRW 405

```

```

RESULT 4
US-09-657-440-18
; Sequence 18, Application US/09657440
; Patent No. 6509455
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/657,440
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/320,878
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: CIP OF 09/141,908
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 18
; LENGTH: 416
; TYPE: PRN

```

```
; ORGANISM: Streptomyces venezuelae
US-09-637-440-18

Query Match      17.0%; Score 37-5; DB 4; Length 416;
Best Local Similarity 28.6%; Pred. No. 7.7e-31;
Matches 98; Conservative 62; Mismatches 162; Indels 21; Gaps 7;

QY 76 EDVRFSSCPFFIPRAGEAYDPIPTSMDEPPEQOROFALANOVGMPPVVDKLENRQELA 135
DB 72 KWR--NSTPTTEAEEAALHNWLES--DEPHTRLRKLVAREFTNRVVELLAPRQEI 127

QY 136 CSLIESL--RPGQCNTFYDAPPPFIRFMLLAGLPEEDIPHLKYLTDQMTGDSMTF 193
DB 128 DGLVDAMLAAPDGRADLMESLAWPLDITV--SELLGVPEPDRAAFRVWTDFAVFPDPAQA 187

QY 194 ABAKALYLYLPIIQRQKXTDAIS--TVANGVNGRPITTSDEAKRMCGLLLVGLDT 252
DB 188 QTAMAMSGYLSRLIDSKRGDGEDLLSALVTSDEGDSRLTSEELGMHAILLVAGHET 247

QY 253 VNFLSFSEFLAKSPHROELIERPELIPAAACEELLRRFSIWDGRIILTSYEF---- 307
DB 248 TYNLIANGYALLSHPDQLAARADMTLLDGAVEMLR-----YEGPVESATYRFPV 302

QY 308 --HGVLKKGQDQILLPQM--SGLDERKNACPMHYDFSRQKVSHTTFGSHGLCGHARR 365
DB 303 DLDGTVIPAGDTVLVWJADAHRTPEPFPDHRFDIRDTAGHLAFGCHGIFCIGAPLARL 362

QY 366 EIVVLKWLTRIPDS--IAPCAQIOHKSIGVSGVQALPLW 406
DB 363 EARIAVALLERCPCJALDVSFGLVWYNPMIRGLKAUPIRW 405

RESULT 5
US-09-252-991A-17836
; Sequence 17836, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc C. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17836
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17836

Query Match      16.3%; Score 355; DB 4; Length 437;
Best Local Similarity 25.8%; Pred. No. 4.8e-23;
Matches 96; Conservative 68; Mismatches 158; Indels 50; Gaps 8;

QY 63 WIATRGQIREAYEDYRHFSSCCFFIPRA-----GEAYDFIPTSMD 104
DB 60 WYVTR-----YRDARKVLNH--PGVRDARQAELYAKRTGTSFRAGIEGLSHHMLNLD 111

QY 105 PPEQORFALANOVGMPPVVDKLENRIQELACSLIESLRPQGCNFTEDYAEPPPIRIFM 164
DB 112 PPDHTLRSLVGAFTRQVERLCQPHRIETELLDMAGREQADLMADFAILTLTAVIF 171

QY 165 LLAGLPEEDIPHLKYLTD---QWTRPDGSMTPFAEAKLYDYLIPIIEQRQKPGTDAIS 221
DB 172 ELLGIPEAREHARQSWERQAELLSPEEAQALADAQ---VDYLRVLLEAKRRQPADVYS 228

QY 222 IVANGVNGRPITTSDEAKRMCGLLLVGLDVTNVFLSFSMEFLAKSPHROELIERPEL 281
DB 229 GLVQAADSGSQSEAEFLVSNHLMISGFTTMMIGNALVTLLVNPQELALRAQFELL 288

; ORGANISM: Streptomyces venezuelae
US-09-637-440-18

Query Match      15.9%; Score 346; DB 6; Length 403;
Best Local Similarity 29.2%; Pred. No. 3.9e-28;
Matches 106; Conservative 63; Mismatches 164; Indels 30; Gaps 13;

QY 63 WIATRGQIREAYEDYRHFSSCC-----PTPREAGEAYDFIPT--SMDPPEQORFALA 115
DB 52 WLVTIRHQDVRAVLGDER--FSADAHRTGFFELTAGGRIIIGTNTPTFLRMDDEPHARLRML 110

QY 116 NOVGMPPVVDKLENRIQELACSLIESLRP--QGQCNTEDYAEPPPIRIFMLLAGLPEEDI 174
DB 111 TADFIYKVEAMSPQVRUADDLVDRMTTGTSTADLVTEFALPLPSLVICLLIGVPYEDH 170

QY 175 PHL----KYLTQMTRPDGSMTFAEAKLYDYLIPIIEQRQKPGTDAIS--IVANGV 229
DB 171 AFPQERSRVLLTLRS--PE---EYRAAQDELLLEYLARLARTKRERPDDAIISRLVARGELD 227

QY 230 GRPITSDEAKRMCGLLLVGLDVTNVFLSFSMEFLAKSPHROELIERPELIPAAACELL 289
DB 228 DTQIAT-----MGRLLLVAGHETTANMTALSTLVLLNPNQLARLRAEPALVKGAVELL 282

QY 290 RRFSLVADG--RLTSDYEFHGVLKKGQDQILLPQMLSGLDERKNACP--MHVDFSRQKV 345
DB 283 RYLTVHNGVPRATIEDVLIGRTIAGEGVLC--MISSANRDAEVPFGDDLDVADAR 340

QY 346 SHTFGSHGLCGHARREIIVTKWLTTRIPDFSIA--PGAQIOHKSIGI--VSGVQALP 403
DB 341 RHVAFGFGVHCIGQPLARVELQIA--ETLLRLRLPLAVPHEEIPRGMAYGVESDP 400

QY 404 LW 406
DB 401 IAW 403

RESULT 7
US-09-385-028-12
; Sequence 12, Application US/09385028
; Patent No. 6232106
; GENERAL INFORMATION:
```

APPLICANT: Susan E. Jensen
APPLICANT: Kwamena A Aidoo
APPLICANT: Ashish S. Paradkar
TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic
Patent No. 6232106
TITLE OF INVENTION: Acid Biosynthesis
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC
STREET: The Jennifer Building, 400 Seventh Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/385,028
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/790,462
FILING DATE: 29-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: D. Douglas Price
REGISTRATION NUMBER: 24,514
REFERENCE/DOCKET NUMBER: 1418/P57452US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638-6666
TELEFAX: (202) 39305350
TELEX: RCA 248593 IDEA UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 409 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-385-028-12

Query Match 15.6%; Score 339; DB 3; Length 409;
Best Local Similarity 29.2%; Pred. No. 2.2e-27;
Matches 123; Conservative 60; Mismatches 184; Indels 54; Gaps 18;
QY 16 PHVPEHLVFDPMYFNLSAGVOEAWVLQESNVPDLVWTRCGH-WIATRGQLIREA 74
DB 14 PAYPMHRVCPVD---PPQLAGLRSKAASRT-----LW---DGSQVWLVTSHAGARAV 62
QY 75 YEDYRHFS-SECFPIP-----REAGAYDFIPTSMDDPPEQORQFRA-----LANQV 118
DB 63 LGSRRFTAVTSAPGFMLTTSQVLVANPESASFI--RMDDPQHSRLSMLTRDFLARRA 120
QY 119 VGM-PVYDKLENRIQELACSLIESLRPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHL 177
DB 121 EALRPVAVREL---LDELGLGVKGERP---VDLVAGLTIPVPSRVITLLFSGADDRREFI 174
QY 178 K-----YLTQDMTPRDSMTFAEAKELYDILPIEIORRQKPGTDAISIVANGQVGRPI 233
DB 175 EDRSAVLIDRGYTP---QVAKARDELQVLRVLEVERIENPGTDLISRLWIDQVRPHL 231
QY 234 TSEAKEMCGLLVGGDLTVNFLSFSMEFLAKSPHEHQELIERPELIPACCELLRRFS 293
DB 232 RVEMVPMCLRLVAGHGTTTSQASLSLLTDPPELAGRLTDPALPXAVELLRFHS 291
QY 294 LVADG---RLTSDYEFHGVOLKGDQILLPQLSGDLERKNACPMVDFSRQKVSHTTFG 351
DB 292 IVQNGLARAARVEDVQDDVLIRAGEGVVLSAGNRDETVEPPDRVDVDRDARRHLAFG 351
QY 352 HGSHLCAGQHAR---REIIVTLKEWLTTRIPDSIA-PGAQIQHKSQIVS-GVOALPLVW 406

DB 352 EEMHQCLGQWLARVELEBILAAVLRW---PGARLAVPPELDFRHEVSSVGLCALPVTW 408
QY 407 D 407
DB 409 Z 409
RESULT 8
US-09-726-614-12
Sequence 12, Application US/09726614
Patent No. 6514735
GENERAL INFORMATION:
APPLICANT: Susan E. Jensen
APPLICANT: Kwamena A Aidoo
APPLICANT: Ashish S. Paradkar
TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic
Patent No. 6514735
TITLE OF INVENTION: Acid Biosynthesis
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC
STREET: The Jennifer Building, 400 Seventh Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/726,614
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/790,462
FILING DATE: 29-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: D. Douglas Price
REGISTRATION NUMBER: 24,514
REFERENCE/DOCKET NUMBER: 1418/P57452US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638-6666
TELEFAX: (202) 39305350
TELEX: RCA 248593 IDEA UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 409 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-726-614-12

Query Match 15.6%; Score 339; DB 4; Length 409;
Best Local Similarity 29.2%; Pred. No. 2.2e-27;
Matches 123; Conservative 60; Mismatches 184; Indels 54; Gaps 18;
QY 16 PHVPEHLVFDPMYFNLSAGVOEAWVLQESNVPDLVWTRCGH-WIATRGQLIREA 74
DB 14 PAYPMHRVCPVD---PPQLAGLRSKAASRT-----LW---DGSQVWLVTSHAGARAV 62
QY 75 YEDYRHFS-SECFPIP-----REAGAYDFIPTSMDDPPEQORQFRA-----LANQV 118
DB 63 LGSRRFTAVTSAPGFMLTTSQVLVANPESASFI--RMDDPQHSRLSMLTRDFLARRA 120
QY 119 VGM-PVYDKLENRIQELACSLIESLRPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHL 177
DB 121 EALRPVAVREL---LDELGLGVKGERP---VDLVAGLTIPVPSRVITLLFSGADDRREFI 174
QY 178 K-----YLTQDMTPRDSMTFAEAKELYDILPIEIORRQKPGTDAISIVANGQVGRPI 233

Db 175 EDSAVLIDRGYTP--QVAKARDELQYLRELVBEIENPGTDLISRLVLDQVRPHGL 231
QY 234 TSDAKRMCGLLVGLDVTWNFLSFMELAKSPHQRQELIERPELIPACBELLARFS 293
Db 232 RVEEMVPMCRLLVAGHGTTSQASLSLLTDPELAGRLTDPALLPKAVBELLARFS 291
QY 294 LVADG--RLITSYEFHGVOLKKGDOILLPQMLSGLDERKNACPMHVDFSRQKVSHTTFG 351
Db 292 IVQNGLARAAVEDVQLDVLIRAGEGVVLSLSAGNRDETVPDPDRVDVDRDARHLAFG 351
QY 352 HGSHLCGLQGHAR---REIIVT-KENLTRIPDFSIA-PGAQIQHKSGIVS-GVQALPLVW 406
Db 352 HGMHQCIGQWLARVELEELAAVLKWM---PGARLAVPFEELDFRHEVSSVGLGALPVTW 408
QY 407 D 407
Db 409 Z 409
RESULT 9
US-09-385-040-12
; Sequence 12, Application US/09385040
; Patent No. 6589775
; GENERAL INFORMATION:
; APPLICANT: Jensen, Susan E
; APPLICANT: Aidoo, Kwamena A
; APPLICANT: Paradkar, Ashish S
; TITLE OF INVENTION: DNA SEQUENCE ENCODING ENZYMES OF CLAVULANIC ACID
; FILE REFERENCE: 09/385,040
; CURRENT FILING DATE: 1999-08-30
; PRIOR APPLICATION NUMBER: US/09/385,040
; PRIOR FILING DATE: 1997-01-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Streptomyces clavuligerus
US-09-385-040-12

Query Match 15.6%; Score 339; DB 4; Length 409;
Best Local Similarity 23.2%; Pred. No. 2.2e-27;
Matches 123; Conservative 60; Mismatches 184; Indels 54; Gaps 18;
QY 16 PHVEHLVDFDFMYNPNLSAGVOEAMAVLQESNVDPDLVTRCNGGH-WIATRGQLIREA 74
Db 14 PAVEMHRVCPVD---PPQLAGLSQKAASRVT---LW---DGSQVWLVTSHAGARAV 62
QY 75 YEDYRHS-SECPFIP-----REAGEAYDFTPTSMOPPEQRFRA-----LANQV 118
Db 63 LGDRRTAVT-SAGPFPMLTRTSQLVRAVNPESASFIR-RMDDFQHSRLRSM-LTRDFLARRA 120
QY 119 VGX--PVYDKLENRIQELACSLIESLRPGCGNFTEDYAEPPFIRFMLLAGLEFEDIPHL 177
Db 121 EALRPVAREL---LDELGLGVKSERP---VDLVAGLTIPVESRWITLLFGAGDREFFI 174
QY 178 K-----YLTQMTREPGSMTFAEAKLYDYLIPIIEQRQRKPGTDAISIVANGQNGRPI 233
Db 175 EDRSAVLIDRGYTFE--QVAKARDELQYLRELVBEIENPGTDLISRLVLDQVRPHGL 231
QY 234 TSDAKRMCGLLVGLDVTWNFLSFMELAKSPHQRQELIERPELIPACBELLARFS 293
Db 232 RVEEMVPMCRLLVAGHGTTSQASLSLLTDPELAGRLTDPALLPKAVBELLARFS 291
QY 294 LVADG--RLITSYEFHGVOLKKGDOILLPQMLSGLDERKNACPMHVDFSRQKVSHTTFG 351
Db 292 IVQNGLARAAVEDVQLDVLIRAGEGVVLSLSAGNRDETVPDPDRVDVDRDARHLAFG 351
QY 352 HGSHLCGLQGHAR---REIIVT-KENLTRIPDFSIA-PGAQIQHKSGIVS-GVQALPLVW 406
Db 352 HGMHQCIGQWLARVELEELAAVLKWM---PGARLAVPFEELDFRHEVSSVGLGALPVTW 408

QY 407 D 407
Db 409 Z 409
RESULT 10
5212296-6
; Patent No. 5212296
; APPLICANT: DEAN, CAROLINE HARDER, PATRICIA A. LETO, KENNETH
; J.; O'KEEFE, DANIEL P.; OMER, CHARLES A.; ROMESSER, JAMES A.
; TEPPERMAN, JAMES M.
; TITLE OF INVENTION: EXPRESSION OF HERBICIDE METABOLIZING
; CYTOCHROMES
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/569,781
; FILING DATE: 23-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 464,499
; FILING DATE: 12-JAN-1990
; APPLICATION NUMBER: 405,605
; FILING DATE: 11-SEP-1989
; SEQ ID NO: 6
; LENGTH: 406
5212296-6

Query Match 15.4%; Score 336; DB 6; Length 406;
Best Local Similarity 26.6%; Pred. No. 4.6e-27;
Matches 105; Conservative 74; Mismatches 192; Indels 24; Gaps 11;
QY 31 PSNLSAGVQ--EAWAVLQESNVDPDLVTRCNGGH-WIATRGQLIREAYEDYR----- 79
Db 17 PSNRSCPQLDPDGYAQLRDTGPHLRHVTLVYGRQAWVTKHEAAKLLAGDPRLSSNRITDD 76
QY 80 HFSSECPFIP--REAGEAYDFTPTSMOPPEQRFALANQVVMVVDKLENRIQELACS 137
Db 77 NFPATSPRFAEVRESPOAF---IGLDPPEHGTTRMTISEFTVKRIKMRPEVEVHVHG 132
QY 138 LIESLRPOS-OCNFTEDYAEPPFIRFMLLAGLEEDIPHLKYLTDQMTDRPGSMTFAEA 196
Db 133 FLDEMLAAGPTADLVSQFALPVPMSVICRLLGVPYADHEFFQDASKRLVQSTDAQSALTA 192
QY 197 KEALYDYLPIPIEORRQKPGTDAL-SIVANGQNGRPTTSDAKRMCGLLVAGGLDVTWN 255
Db 193 RNDLAGYLDGLITQFQTEPGAGLVGALVADQLANGE-IDRSELSTAMLLLAGHETTAS 251
QY 256 FLFSMEFLAKSPHQRQELIERPELIPACBELLARFSL--VADGRILTSYEFHGVOLK 313
Db 252 MTSLSVITLLDHPCEYAAALRADRSILVPGAVEBELLRYLAIADIAGGRVATADIEVEGHLR 311
QY 314 KGDQILLPQMLSGLDERKNACPMHVDFSRQKVSHTTFGHSGLCLGOLARREIIVTLKE 373
Db 312 AGEGVVWVNSIANRDGTVYEDPDALDIHRSARHHLAFGFGVHQCLQONLARLELEVILNA 371
QY 374 WLTRIPDFSIA-PGAQIQHKSG-IVSGVQALPLVW 406
Db 372 LMDRVTPLRLAVPQVLVLRPGTTLIQSNELPVTW 406
RESULT 11
US-08-765-907A-10
; Sequence 10, Application US/08765907A
; Patent No. 6352839
; GENERAL INFORMATION:
; APPLICANT: BLANC, Veronique
; APPLICANT: THIBAUT, Denis
; APPLICANT: BAMAIS-JACQUES, Nathalie
; APPLICANT: BLANCHER, Francis
; APPLICANT: COUZET, Joel
; APPLICANT: BARRIERE, Jean-Claude
; APPLICANT: DEBUSSCHE, Laurent
; APPLICANT: FAMECHON, Alain

APPLICANT: PARIS, Jean-Marc
APPLICANT: EURUC-ROSSET, Gilles
TITLE OF INVENTION: Streptogramins And Method For Preparing Same By
TITLE OF INVENTION: Mutasynthesis
FILE REFERENCE: Streptogramin genes
CURRENT APPLICATION NUMBER: US/08/765,907A
CURRENT FILING DATE: 1997-03-20
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patent in Ver. 2.0
SEQ ID NO 10
TYPE: PRT
ORGANISM: Streptomyces pristinaespiralis
US-08-765-907A-10

Query Match: 14.8%; Score 323.5; DB 4; Length 399;
Best Local Similarity 29.0%; Pred. No. 9.6e-26;
Matches 106; Conservative 59; Mismatches 159; Indels 41; Gaps 12;

QY 74 AYEDYRH-----PSSECFIPREAGEAYDRIPTSMDPPEQRFALANQVGM 121
DB 36 AHWFRADVLTASDPGVSSQSLRPPGQALSEQILSVDPDMHTLRRLVSQAFTP 95
QY 122 PVVDKLENR-CELACSLIESLRPOGC-NFTEDYAEFPPIRIFMLLAGLPBEDIPHLKYL 180
DB 96 RTVADLEPRVTELQALLDAV--DGDIFLVADFAFPLPVIVIAELLGVPPADRTLFSW 153
QY 181 TDQMT-----PGSMTFEAKAALYDYLPIIEOROKPGTDA-SIVA 224
DB 154 SDRMLQWQADPADMQFGDDADEYQRLVKEPMRANHAYLHDHVTDRARPANL-SALV 213
QY 225 NGQVNGRPITSDEAKRMCGLLVGLLTVNLFSSMEFLAKSPEHQELIERPE--LIP 282
DB 214 AARVEGRUTDEQIVEGALCIMAGHVSMTLGNVLCLKDHP--RAEAAARADRSLLP 271
QY 283 AACBELLR-RFLSLVADGRILTSDEYFHGVOLKKGQDILLPQVLS-GLDERKNACPMHVD 340
DB 272 ALIEVLRRLPPIITMARVTTKTDTVLAGTTIPAG-RMVVPSLLSANHDEQVETDPDHL 330
QY 341 SRQKVSHTFHGSHLCLGQHLARREIIVTLKWLTRIPDSIAPGAIQ-HKSGISGV 399
DB 331 AREG-RQIAFGHGHYCLGASLARLEGRALAEALFDREPFDSPTDGAKLRYHRDGLF-GV 388
QY 400 QALPL 404
DB 389 KNLPL 393

RESULT 12
US-08-102-863-11
Sequence 11, Application US/08102863
Patent No. 5466590
GENERAL INFORMATION:
APPLICANT: SARIASLANI, SIMA
TITLE OF INVENTION: CONSTITUTIVE
TITLE OF INVENTION: EXPRESSION OF P450SOY
TITLE OF INVENTION: AND FERREDOXIN-SOY IN
TITLE OF INVENTION: STREPTOMYCES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS
ADDRESSEE: AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: USA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0,
SOFTWARE: Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/102,863
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/807,001
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: GALLEGOS, R. THOMAS
REGISTRATION NUMBER: 32,692
REFERENCE/DOCKET NUMBER: CR-9000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-7342
TELEFAX: 302-892-7949
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 412 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-102-863-11

Query Match 14.5%; Score 315; DB 1; Length 412;
Best Local Similarity 27.2%; Pred. No. 8.2e-25;
Matches 84; Conservative 56; Mismatches 163; Indels 6; Gaps 5;

QY 103 MDPPEQRFALANQVGMVVDKLENRIOELACSLIESLRPOG-QCNFTEDYAEFPPIR 161
DB 105 VDDPEHTQTRMLIPTFSVKRIGALRPRIQETVVDLDMERQGPALVSAFALPVFSM 164
QY 162 IFMLLAGLPBEDIPHLKYLTDQMTRPDGSMTFAEAKAALYDYLPIIEOROKPGTDAIS 221
DB 165 VICALLGVYADHAFFERSQRLLRGCGADDVNARDEEYLGALIDRKRAEPGDLID 224
QY 222 IVANGVNGRPITSDAEAKRMCGLLVGLLTVNLFSSMEFLAKSPEHQELIERPELI 281
DB 225 ELIHRDHPDGVDRQELVAFVILLIAGHETTANNISLGTFTLLSHPEQLAALRAGGTST 284
QY 282 PAACELLRRFLSVADG--RLTSDYEPHGVOLKKGDDOILLPQMLSGLDERKNACPMHVD 339
DB 285 AVVUELL-RFLSLVADGRILTSDEYFHGVOLKKGDDOILLPQMLSGLDERKNACPMHVD 343
QY 340 FSROKVSHTFHGSHLCLGQHLARREIIVTLKWLTRIPDPSIA-PGAIQHKSG-IVS 397
DB 344 WDRPARHHLAFGFGVHQCLQGMARAEIDIAKTLFERLPGRLAVPAHEIRHKPGDTIQ 403
QY 398 GVQALPLVW 406
DB 404 GLLDLPVAV 412

RESULT 13
PCT-US92-10885-11
Sequence 11, Application PC/TUS9210885
GENERAL INFORMATION:
APPLICANT: SARIASLANI, SIMA
TITLE OF INVENTION: CONSTITUTIVE
TITLE OF INVENTION: EXPRESSION OF P450SOY
TITLE OF INVENTION: AND FERREDOXIN-SOY IN
TITLE OF INVENTION: STREPTOMYCES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS
ADDRESSEE: AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: USA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch,
MEDIUM TYPE: 1.0 MB

```

; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh System, 6.0
; SOFTWARE: Microsoft Word, 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10885
; FILING DATE: 199212-6
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: GALLEGOS, R. THOMAS
; REGISTRATION NUMBER: 32,692
; REFERENCE/DOCKET NUMBER: CR-9000-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-7342
; TELEFAX: 302-892-7949
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 412 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; PCT-US92-10885-11

Query Match      14.5%; Score 315; DB 5; Length 412;
Best Local Similarity 27.2%; Pred. No. 8.2e-25;
Matches 84; Conservative 56; Mismatches 163; Indels 6; Gaps 5;

QY 103 MDPPEQRFALANQVGMVVVKLENRIQELACSLIESLRPQG-OCNFTEDYAEPPPIR 161
DB 105 VDDPHNTQRMPLTFSVGRIGALRPRTQETVDRLDDAMERQGPFAELVSAPALPVP 164
QY 162 IFMLLAGIPEDIPHLKYLTQMTPEPGSMTFAEAKELALDYLIPIISORRQKPGCTAIS 221
DB 165 VICALLGVYADHAFPEERSORLLRPGADENVNRARDEEYLGALIDRKRAEPGGGLJ 224
QY 222 IVANGVNGRPITTSDEAKRMGLLVGGGLTVVNFSLFSMEFLAKSPHQRQELIERPELI 281
DB 225 ELIHRDHPDGPVDRQLAVAFVILLIAGHETTANNISLGLTFLLSHPQLAALRAGGTST 284
QY 282 PAACEELLRRSLVADG--RLTSDYEFHGVOLKGGDQILLPQMLSGLDERKNAKCPMHVD 339
DB 285 AVVVEELL-RLSLAEGQLRLATEDMEVDGATIRKGGVVFSTSLINRDADVFPRAETLD 343
QY 340 FSRQKVSHTTGCHSHLCLGQHLARREIIVTLKEMLTTRIPDSIA-PGNAQIOKSG-IVS 397
DB 344 WDRPARHHLAFGFGVHQCLGQKLARAEIDIAKRTLFFRLPGURLAVPAHEIRKPGDTIQ 403
QY 398 GWQALPLVW 406
DB 404 GLLDLPVAV 412

RESULT 14
US-09-335-409-8
; Sequence 8, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335,409
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Sorangium cellulosum

Query Match      12.7%; Score 277.5; DB 3; Length 419;
Best Local Similarity 24.0%; Pred. No. 8.4e-21;
Matches 101; Conservative 67; Mismatches 178; Indels 75; Gaps 14;

US-09-335-409-8
Query Match      12.7%; Score 277.5; DB 3; Length 419;
Best Local Similarity 24.0%; Pred. No. 8.4e-21;
Matches 101; Conservative 67; Mismatches 178; Indels 75; Gaps 14;

QY 6 OSNANLAPLPHVPRHVLVDFDMYNFSLSAGVQE-----AWAVLOESNVPDLVWTRCNGG 61
DB 3 QEQANQSETKP-----AFDFKFPAP-----GYADFFPAIERLREA-TPIFYWD--EGR 48
QY 62 HWIATRGQ-----LIREAYDYRHFSSCEPFIIPREAGEAYDT-PTSMOPPEQR 109
DB 49 SWLTRYHDVSAVFRDERFAVSKEEWSAEYSSAIP-----ELSDMKKXGLFGLPPEDHA 104
QY 110 QERALANQVGMVVVKLENRIQELACSLIESLRPQGCNFTEDYAEPPPIRIFMLAGL 163
DB 105 KVRKLVNPSFTSRAIDLLRAEIQRTVDQLLDARSQGEEDFVVRDYAEGLPKRAISALLKV 164
QY 170 PEEDIPHLKYLTDQMTDPGSMTFAEAKELALDYLIPI-----I 207
DB 165 PAE-----CDEKERRFGSAT-----ARALGVGLVQVDEETKTLVASVTGCLALLHDV 212
QY 208 IEQREQKP-GTDAISIVANGVNGRPITTSDEAKRMGLLVGGGLTVVNFSLFSMEFLAK 266
DB 213 LDERRRNPLENDVLTWLLQAEADGSRSLSKELVALVGAIIAAGTDTTIVLIAFAVLNLLR 272
QY 267 SPEHQRQELIERELLIPAAACEELLRRFSLVADG--RLTSDYEFHGVOLKGGDOI--LLPQ 322
DB 273 SPEALVKAEPLGRNALDEVLRFDNLIRIGTFPARQDLEYCGASIKKGENVFLIPS 332
QY 323 MLSGGLDERKNAKCPMEVDFSRQKVSHTTGCHSHLCLGQHLARREIIVTLKEMLTTRIPDS 382
DB 333 ALR--CGTVFSRPDVEDVVRDTCASLAYGRGHVCDGVSRLARLEAEIANGVTIFRRFP 390
QY 383 I 383
DB 391 L 391

RESULT 15
US-09-413-814-71
; Sequence 71, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bloecker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hofle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
; FILE REFERENCE: PCT/US 99/23535
; CURRENT APPLICATION NUMBER: US/09/413,814
; CURRENT FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; EARLIER FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 71
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Sorangium cellulosum

US-09-413-814-71
Query Match      12.7%; Score 277.5; DB 3; Length 419;
Best Local Similarity 24.0%; Pred. No. 8.4e-21;
Matches 101; Conservative 67; Mismatches 178; Indels 75; Gaps 14;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 18:53:57 ; Search time 34.974 Seconds
(without alignments)
3108.683 Million cell updates/sec

Title: US-09-246-451A-13

Perfect score: 2177

Sequence: 1 TTTTSSNNANLAPLPHVFE.....IVSGVQALPLVNDPATYKAV 414

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 262633353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA: *

```

1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2177	100.0	414	15	US-10-453-104-13
2	2169	99.6	414	15	US-10-453-104-11
3	2165	99.4	414	15	US-10-453-104-2
4	2162	99.3	414	15	US-10-453-104-12
5	372	17.1	399	14	US-10-314-657-37
6	366.5	16.8	404	14	US-10-214-446-50
7	365.5	16.8	404	14	US-10-214-446-40
8	365	16.8	416	14	US-10-156-761-14937
9	363.5	16.7	416	9	US-09-861-289-39
10	363.5	16.7	416	10	US-09-860-846-39
11	363.5	16.7	416	10	US-09-888-3848-39
12	363.5	16.7	416	10	US-09-836-821-39
13	363.5	16.7	416	10	US-09-793-778-18
14	363.5	16.7	416	14	US-10-201-365-13
15	363.5	16.7	416	14	US-10-160-539-18

```

16 363.5 16.7 416 14 US-10-271-889-39
17 357 16.4 405 14 US-10-214-446-38
18 356 16.4 393 14 US-10-156-761-9525
19 354 16.3 425 14 US-10-214-446-20
20 352.5 16.2 399 14 US-10-156-761-9914
21 352.5 16.2 411 14 US-10-156-761-8376
22 348 16.0 418 12 US-10-389-647-559
23 343.5 15.8 457 14 US-10-156-761-11073
24 339.5 15.6 388 14 US-10-156-761-13776
25 337.5 15.2 408 14 US-10-214-446-4
26 330 15.2 392 14 US-10-214-446-32
27 326.5 15.0 405 14 US-10-214-446-2
28 323 14.8 401 14 US-10-156-761-8710
29 323 14.8 421 14 US-10-156-761-9703
30 323 14.8 428 14 US-10-201-213-6
31 322 14.8 409 15 US-10-458-201-12
32 319 14.7 404 14 US-10-214-446-16
33 319 14.7 418 14 US-10-214-446-22
34 317.5 14.6 395 14 US-10-214-446-14
35 315.5 14.5 404 14 US-10-156-761-14659
36 307 14.1 421 14 US-10-214-446-18
37 304 14.0 412 14 US-10-214-446-36
38 303.5 13.9 404 14 US-10-156-761-10431
39 300.5 13.8 415 14 US-10-214-446-46
40 298 13.7 475 14 US-10-145-415-22
41 297.5 13.7 402 14 US-10-305-032-8
42 297 13.6 470 14 US-10-145-415-6
43 296.5 13.6 430 9 US-09-718-826-4117
44 296 13.6 399 14 US-10-156-761-7959
45 295 13.6 429 14 US-10-145-415-14

```

ALIGNMENTS

```

RESULT 1
US-10-453-104-13
; Sequence 13, Application US/10453104
; Publication No. US20030207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Hyun Joo
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
; FILE REFERENCE: 4058/1E827-US3
; CURRENT APPLICATION NUMBER: US/10/453,104
; PRIOR FILING DATE: 2001-06-02
; PRIOR APPLICATION NUMBER: US 09/661,093
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 09/246,451
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/086,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/106,834
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant M7-8H
US-10-453-104-13

```

Query Match 100.0%; Score 2177; DB 15; Length 414;
Best Local Similarity 100.0%; Pred. NO. 9.8e-211;
Matches 414; Conservative 0; Mismatches 0; Gaps 0;

QY 1 TTETIQSNANLAPLPPHVPHELVDFDXYNPSNLSAGVQEAQAVLQESNVDPDLVWTRCNG 60
DB 1 TTETIQSNANLAPLPPHVPHELVDFDXYNPSNLSAGVQEAQAVLQESNVDPDLVWTRCNG 60
QY 61 GHWIATRGQLIREAYEDYRHFSECFPIPREAGEAYDFIPTSDPPEQRCFRALANQVVG 120
DB 61 GHWIATRGQLIREAYEDYRHFSECFPIPREAGEAYDFIPTSDPPEQRCFRALANQVVG 120
QY 121 MPVVDKLENRIQELACSLIESLRPOQCNTFYAEPPPIRIFMLLAGLPEEDIPHLKYL 180
DB 121 MPVVDKLENRIQELACSLIESLRPOQCNTFYAEPPPIRIFMLLAGLPEEDIPHLKYL 180
QY 181 TDQMTRPDGSMTFAEAKAALYDYLPIIEORROKPGTDAISIVANGQVNGRPITSDAKR 240
DB 181 TDQMTRPDGSMTFAEAKAALYDYLPIIEORROKPGTDAISIVANGQVNGRPITSDAKR 240
QY 241 MFGLLVGGGLDVTNNFLSFSMEFLAKSPEHRQELIERPERIPAAACBELLRRFSLVADGRI 300
DB 241 MFGLLVGGGLDVTNNFLSFSMEFLAKSPEHRQELIERPERIPAAACBELLRRFSLVADGRI 300
QY 301 LTSDFYFHGVQKKGQOILLPOMLSGLDERKNACPMHVDFSROKVSHTTFHGSHLCLGQ 360
DB 301 LTSDFYFHGVQKKGQOILLPOMLSGLDERKNACPMHVDFSROKVSHTTFHGSHLCLGQ 360
QY 361 HLARREIIVTLKWLTRIPDPSIAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414
DB 361 HLARREIIVTLKWLTRIPDPSIAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414

RESULT 2
US-10-453-104-11
; Sequence 11, Application JS/10453104
; Publication No. US20030207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Hyun Joo
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
; FILE REFERENCE: 4058/1E827-US3
; CURRENT APPLICATION NUMBER: US/10/453,104
; CURRENT FILING DATE: 2003-06-02
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 09/246,451
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/086,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/106,834
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant M7-4H
US-10-453-104-11

Query Match 99.6%; Score 2169; DB 15; Length 414;
Best Local Similarity 99.8%; Pred. No. 6.3e-210;
Matches 413; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TTETIQSNANLAPLPPHVPHELVDFDXYNPSNLSAGVQEAQAVLQESNVDPDLVWTRCNG 60
DB 1 TTETIQSNANLAPLPPHVPHELVDFDXYNPSNLSAGVQEAQAVLQESNVDPDLVWTRCNG 60
QY 61 GHWIATRGQLIREAYEDYRHFSECFPIPREAGEAYDFIPTSDPPEQRCFRALANQVVG 120

DB 61 GHWIATRGQLIREAYEDYRHFSECFPIPREAGEAYDFIPTSDPPEQRCFRALANQVVG 120
QY 121 MPVVDKLENRIQELACSLIESLRPOQCNTFYAEPPPIRIFMLLAGLPEEDIPHLKYL 180
DB 121 MPVVDKLENRIQELACSLIESLRPOQCNTFYAEPPPIRIFMLLAGLPEEDIPHLKYL 180
QY 181 TDQMTRPDGSMTFAEAKAALYDYLPIIEORROKPGTDAISIVANGQVNGRPITSDAKR 240
DB 181 TDQMTRPDGSMTFAEAKAALYDYLPIIEORROKPGTDAISIVANGQVNGRPITSDAKR 240
QY 241 MFGLLVGGGLDVTNNFLSFSMEFLAKSPEHRQELIERPERIPAAACBELLRRFSLVADGRI 300
DB 241 MFGLLVGGGLDVTNNFLSFSMEFLAKSPEHRQELIERPERIPAAACBELLRRFSLVADGRI 300
QY 301 LTSDFYFHGVQKKGQOILLPOMLSGLDERKNACPMHVDFSROKVSHTTFHGSHLCLGQ 360
DB 301 LTSDFYFHGVQKKGQOILLPOMLSGLDERKNACPMHVDFSROKVSHTTFHGSHLCLGQ 360
QY 361 HLARREIIVTLKWLTRIPDPSIAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414
DB 361 HLARREIIVTLKWLTRIPDPSIAPGAQIOHKSGIVSGVQALPLVWDPATTKAV 414

RESULT 3
US-10-453-104-2
; Sequence 2, Application US/10453104
; Publication No. US20030207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Hyun Joo
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
; FILE REFERENCE: 4058/1E827-US3
; CURRENT APPLICATION NUMBER: US/10/453,104
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 09/661,093
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 09/246,451
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/086,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/106,834
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 414
; TYPE: PRT
; ORGANISM: P. Putida
US-10-453-104-2

Query Match 99.4%; Score 2165; DB 15; Length 414;
Best Local Similarity 99.5%; Pred. No. 1.6e-209;
Matches 412; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 TTETIQSNANLAPLPPHVPHELVDFDXYNPSNLSAGVQEAQAVLQESNVDPDLVWTRCNG 60
DB 1 TTETIQSNANLAPLPPHVPHELVDFDXYNPSNLSAGVQEAQAVLQESNVDPDLVWTRCNG 60
QY 61 GHWIATRGQLIREAYEDYRHFSECFPIPREAGEAYDFIPTSDPPEQRCFRALANQVVG 120
DB 61 GHWIATRGQLIREAYEDYRHFSECFPIPREAGEAYDFIPTSDPPEQRCFRALANQVVG 120
QY 121 MPVVDKLENRIQELACSLIESLRPOQCNTFYAEPPPIRIFMLLAGLPEEDIPHLKYL 180
DB 121 MPVVDKLENRIQELACSLIESLRPOQCNTFYAEPPPIRIFMLLAGLPEEDIPHLKYL 180
QY 181 TDQMTRPDGSMTFAEAKAALYDYLPIIEORROKPGTDAISIVANGQVNGRPITSDAKR 240

Db 181 TDQMPDGSMTFAEAKAALDYLIPIIEQRQKPGTDALSIIVANGQVNGRPITSDEAKR 240
QY 241 MFCLLVGLDVTVNF-SFSMEFLAKSPERHQLIERPERIPACBELLRFSIVADGRI 300
Db 241 MFCLLVGLDVTVNF-SFSMEFLAKSPERHQLIERPERIPACBELLRFSIVADGRI 300
QY 301 LTSDEYEHGVLKGGQILLPQMLSGLDERKNACPMHVDFSRQKVSHTTTFGHSHLCGQ 360
Db 301 LTSDEYEHGVLKGGQILLPQMLSGLDERKNACPMHVDFSRQKVSHTTTFGHSHLCGQ 360
QY 361 HLARREIIVTLKEMWTRIPDPSIAPGAQIOHKSGIVSGVQALPLVMDPATTKAV 414
Db 361 HLARREIIVTLKEMWTRIPDPSIAPGAQIOHKSGIVSGVQALPLVMDPATTKAV 414

RESULT 4

US-10-453-104-12
; Sequence 12, Application US/10453104
; Publication No. US20030207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Hyun Joo
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
; FILE REFERENCE: 4058/1827-US3
; CURRENT APPLICATION NUMBER: US/10/453,104
; PRIOR FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 09/661,093
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 09/246,451
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/086,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/106,834
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 4-4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant M7-6H
US-10-453-104-12

Query Match 99.3%; Score 2162; DB 15; Length 414;
Best Local Similarity 99.5%; Pred. No. 3.2e-209;
Matches 412; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TLETIQSNANLAPLPHPVHELVDYFDMYFNSLSAGVQEAAYLQESNVPDLVWTRCNG 60
Db 1 TLETIQSNANLAPLPHPVHELVDYFDMYFNSLSAGVQEAAYLQESNVPDLVWTRCNG 60
QY 61 GHWIATRGQIREAYEDYRHFSSECPITPRAGAYDFIPTSDMPPEQORFRALANQVVG 120
Db 61 GHWIATRGQIREAYEDYRHFSSECPITPRAGAYDFIPTSDMPPEQORFRALANQVVG 120
QY 121 YEVVDKLENRIQELACSLIESLRPGQCNFTEDYAEPPPIRIFMLLAGLPEDIPHLKYL 180
Db 121 MPVVDKLENRIQELACSLIESLRPGQCNFTEDYAEPPPIRIFMLLAGLPEDIPHLKYL 180
QY 181 TDQMPDGSMTFAEAKAALDYLIPIIEQRQKPGTDALSIIVANGQVNGRPITSDEAKR 240
Db 181 TDQMPDGSMTFAEAKAALDYLIPIIEQRQKPGTDALSIIVANGQVNGRPITSDEAKR 240
QY 241 MFCLLVGLDVTVNF-SFSMEFLAKSPERHQLIERPERIPACBELLRFSIVADGRI 300
Db 241 MFCLLVGLDVTVNF-SFSMEFLAKSPERHQLIERPERIPACBELLRFSIVADGRI 300

QY 301 LTSDEYEHGVLKGGQILLPQMLSGLDERKNACPMHVDFSRQKVSHTTTFGHSHLCGQ 360
Db 301 LTSDEYEHGVLKGGQILLPQMLSGLDERKNACPMHVDFSRQKVSHTTTFGHSHLCGQ 360
QY 361 HLARREIIVTLKEMWTRIPDPSIAPGAQIOHKSGIVSGVQALPLVMDPATTKAV 414
Db 361 HLARREIIVTLKEMWTRIPDPSIAPGAQIOHKSGIVSGVQALPLVMDPATTKAV 414
RESULT 5
US-10-314-657-37
; Sequence 37, Application US/10314657
; Publication No. US20030175889A1
; GENERAL INFORMATION:
; APPLICANT: SHEN, Ben
; APPLICANT: CHENG, Yi-Qiang
; APPLICANT: TANG, Gong-Li
; TITLE OF INVENTION: Discrete Acyltransferases Associated with Type I Polyketide
; TITLE OF INVENTION: Syntheses and Methods of Use
; FILE REFERENCE: 054030-0021
; CURRENT APPLICATION NUMBER: US/10/314,657
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: PCT/US02/08937
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 60/278,935
; PRIOR FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 37
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Streptomyces atroolivaceus
US-10-314-657-37

Query Match 17.1%; Score 372; DB 14; Length 399;
Best Local Similarity 28.1%; Pred. No. 1.8e-28;
Matches 110; Conservative 70; Mismatches 176; Indels 36; Gaps 9;

QY 38 VQEAWAVLOESNVPDLVWTRCNGGHWIATRGQIREAYEDYRHFSSECPITPREAGE--- 94
Db 21 IHPKFAELRETDPLARVLPYGGEGWVTR-----YDDVRAANSDFRFSQAQIGEDTF 73
QY 95 -----AYDFIPTSDMPPEQORFRALANQVVGMPVDKLENRIQELACSLIESLRPQ 147
Db 74 RTTPLARRSDTI-LSLDPPEHTRRLRLSKAFTARRMGAMQSWLEELFAGLIDGVERTGH 132
QY 148 -CNFTEDYAEPPPIRIFMLLAGLPEDIPHLKYLTDQMPDGSMTFAEAK-----A 199
Db 133 PADIVRLAQPTTAVICRLGVPEYDGRFQHWSEVI-----MSTTAYSKEEAVSADMS 187
QY 200 LYDYLIPITIEQRQKPGTDALSIIVANGQVNGRPITSDEAKRMFGLLWGLDVTVNFSLF 259
Db 188 IRAYLADLVASARRAAPHDDLLGLVLSARDDDRLTEDELITFGVTLVAGHETSALQNG 247
QY 260 SWEFLAKSPERHQLIERPERIPACBELLRPSL---VADGRILTSDFEHGVLKGGD 316
Db 248 MVYALLTHEDQSLIREQPELLPRAVEELLRFVPLGNGVGNARIALEDELVELSGGTVRAGE 307
QY 317 QILLPQMLSGLDERKNACPMHVDFSRQKVSHTTTFGHSHLCGQLARREILVTLKEMLT 376
Db 308 GVAAAAVNANDPRAFDPPDRLDITREKNPHLAGHGAHYCLGACLARVELKVAIGELLE 367
QY 377 RIPDPSIA-PCGAQIOHKS-GIVSGVQALPLVW 406
Db 368 RFGRLAVPADQVEMTKGGIFRGFQRLFIAM 399

RESULT 6

US-10-214-446-50
; Sequence 50, Application US/10214446
; Publication No. US20030180742A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, David

```

; APPLICANT: Burk, Mark J.
; APPLICANT: Hitchman, Tim
; APPLICANT: Pujol, Catherine
; APPLICANT: Richardson, Toby
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: P450 ENZYMES, NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: THEM AND METHODS OF MAKING AND USING THEM
; FILE REFERENCE: 09010-500001
; CURRENT APPLICATION NUMBER: US/10/214,446
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 60/309,497
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Bacterial
US-10-214-446-40

Query Match      16.8%; Score 365.5; DB 14; Length 404;
Best Local Similarity 29.8%; Pred. No. 8.1e-28;
Matches 96; Conservative 64; Mismatches 158; Indels 33; Gaps 5;

QY 60 GGHWATRGQIRREAYEDYRHFSSECPPIPREAGEAYDFIPTSM-----103
Db 32 GRTWFLPHHADIRALDERFSAS-----RKAGGFVNFPAEVRPEAFNEALSRIW 85

QY 104 --DPPEQRQRALANQVGMVVDKLENRIQELACSLIESLRPQCCNFTEDYAEPPPIR 161
Db 86 LHDQPEHRQLRQMQCGFTRLITITMEPKIQRCVDELDIAFVXRGSTEFMEYAHPPAK 145

QY 162 IFMLLAGLPEDIEHLKYLTDQMRPDGSM-----TFRAAKE---ALVDYLIPILPEQRQ 213
Db 146 VIAEMLVNPDYFAFVVMSEDLNFAGSLRPTLEMFRAAQDGLLMMQDYFARLLPERE 205

QY 214 KPGTDA-SIVANGVNGRPITSDAKRMFGLLVGLLTVNVLFSMEFLAKSPEHRQE 273
Db 206 NPGDLSVLLSAESEGEWMTAEQVLANCTQIIVAGHETTRNLVANGVELLRYPEQAL 265

QY 274 LIERPERIPACEELLRRFS-LVMDGRILTDYEPHGVLKKGQDQILLPOMLSGLDERN 332
Db 266 LESRPELMPSAVREIMRPESPLQFIRRVAREDFEFGGAEVREGDGLVLMGSAANDPEAF 325

QY 333 ACPMHVDPESQKVGSHTTGHGSHLCLGQHLARE--IVTLKEWLTIRPDPSI 383
Db 326 DDDPTDFLTRNPTGHLAFGWGPHVCVGAALAELEGQVSEFRILLDRPLGLEL 376

RESULT 7
US-10-214-446-40
; Sequence 40, Application US/10214446
; Publication No. US20030180742AL
; GENERAL INFORMATION:
; APPLICANT: Weiner, David
; APPLICANT: Burk, Mark J.
; APPLICANT: Hitchman, Tim
; APPLICANT: Pujol, Catherine
; APPLICANT: Richardson, Toby
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: P450 ENZYMES, NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: THEM AND METHODS OF MAKING AND USING THEM
; FILE REFERENCE: 09010-500001
; CURRENT APPLICATION NUMBER: US/10/214,446
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 60/309,497
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Bacterial

US-10-214-446-40
Query Match      16.8%; Score 365.5; DB 14; Length 404;
Best Local Similarity 29.8%; Pred. No. 8.1e-28;
Matches 118; Conservative 58; Mismatches 189; Indels 31; Gaps 10;

QY 31 PSNLSAGVQBAWVLQESNVPDLVWTRCNGGH--WATRGQIRREAYEDYRHFSSECPPI 88
Db 20 PFNEADGISLADAYEAREEQPLLRVRMAYGEPAWLATR-----YADARLVLGDRRFS 72

QY 89 -----PREAGEAYDFIPTSDMPPEQRQRALANQVGMVVDKLENRIQELACSLI 139
Db 73 RAEGARHDEPRQSEGRDSDGLSMDDPDHTRLRITVAKAFTMHQVEKLRPAVRELLADELI 132

QY 140 ESLRPOG-QCNFTEDYAEPPPIRIFMLLAGLPEDIEHLKYLTDQMRPDGSMT---FAE 195
Db 133 DQWATGAPVDLVEEFALPVPVGVICQLLVPEVDRPRFRANSW-AALSTSSLTAEEDFA 191

QY 196 AKEALYDYLIPILIEQRKQPGTDAISIVANGVNGRPITSDAKRMFGLLVGLLTVN 255
Db 192 NQEEELRAYMGLIEDHRARPREDLITGLIEARDRDRDLTEQELVDLCVGLVAGHETTAT 251

QY 256 FLFSMEFLAKSPEHRQELIERPERIPACEELLRRFSIVADG-----RLTSDYSEHGVO 311
Db 252 QIPNFVVTLLDRPQMNRLRECPPELVPTAVBELM-RFVPLSGSGAFPRATEDVVEGZL 310

QY 312 LKKGDOILLPOMLSGLDERKNACPMHVDPESQKVSHTTFFGSSHLCLGQHLARREIIVTL 371
Db 311 VRAGEPVLVAGAAANDPARFDAQELDLAREGNQHLGFGHGVHCLGAPLARLELEAL 370

QY 372 KEWLTIRPDPSIAGCAQIOHKSQ-IVSGVQALPIVW 406
Db 371 GALLRLRLGLRIA--GDIEWKTQMLVRGPRTLIPVGM 404

RESULT 8
US-10-156-761-14997
; Sequence 14997, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14997
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14997

Query Match      16.8%; Score 365; DB 14; Length 416;
Best Local Similarity 29.1%; Pred. No. 9.5e-28;
Matches 104; Conservative 56; Mismatches 181; Indels 16; Gaps 8;

QY 63 WATRGQIRREAYEDYRHFSSECPPIPR--EAGEAYDFIPT-----SMDPEQRQFRAL 114
Db 63 WYVTGHAARALLSDQRLSSDRT--LPRFPATTEFAVTRRVALLGVDYDDPHEFQRM 120

QY 115 ANQVVGMPVVDKLENRIQELACSLIESLRPQG-QCNFTEDYAEPPPIRIFMLLAGLPED 173
Db 121 LVESFTLKGAALRPRIOETVDGLLGAQPPAEALQPPAEALPLPSMVICALLGVPYAD 180
```

```
QY 174 IPHLYTQXTRPDGSMWTFEAEKALYDYLPIIEQRKQKGTDAISIVANGQVNGRPI 233
Db 181 HDPFESQSRLLRGFGIAEAGDADADYALIDRKRKEPGDGLLDLIQEQLNRGT 240
QY 234 TSDAKYFGLLVGGDTVNFISFSEFLAKSPEHQELIERPERIPAAACEILLRFS 293
Db 241 DRALVSAJLLAGHETTANMISLSTFTLRPEQLAEIRABEPGLMPAAVEILL-RFL 299
QY 294 JVARG--RIITSDYEFHGVQKGDQILLPQLSGDERKNACPMHVDFFSKVSHHTFG 351
Db 300 SIADGLLRVATEDIEVAGTT--RADEGVVFATSVINRDAAGFAEPDADLWHSARHVAFG 359
QY 352 HSHLCLGQHLARERIIIVTKWETRIPOFSI-APGAQIQKSG-IVSGVQALPLVW 406
Db 360 FGHQCLGONLARAEMETALGTLFERLPGLRLAAPADEIPFKGTTIQGMLELPTVW 416
```

RESULT 9

```
US-09-861-289-39
; Sequence 39, Application US/09861289
; Patent No. US2002010897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/861,289
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRP
; ORGANISM: Streptomyces venezuelae
US-09-861-289-39
```

```
Query Match 16.7%; Score 363.5; DB 9; Length 416;
Best Local Similarity 28.3%; Pred. No. 1.3e-27;
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;
```

```
QY 76 EDYRHSSECPPIPREAGEAYDFTSMDDPPQROFRALANQVGMVVDKLENRIQELA 135
Db 72 KDW--NSTPLTEAEALNHNWLES--DPPRHLRLKLVAREFTMRRVVELLRPRVQEI 127
QY 136 CSLIESL--RPOGQCNFTEDYAEPPPIEFMLLAGLPEEDIHKLXLTQMTDPPGSM 193
Db 128 DGLVDAMLAAPDGRADLMESLAWPLITWISLGVPEPDRAAFRWTDVAFVDDPAQA 187
QY 194 AEAKALYDYLPIIEQRKQKGTDAIS-IVANGQVNGRPIITSDEAKRMFGLLVGG 252
Db 188 QTAAEMSYLSRLIDSKRGQDGLLSALVTSDEGSRLLSEELGMAHILLVAGHET 247
QY 253 VVNF--SFMSEFLAKSPEHQELIERPERIPAAACEILLRFSIVADGRILTSYEF 307
Db 248 TYNLIANGMYALLSHPDQALRADMTLLDGAVEEMLR-----YEGPVESATYRFP 302
QY 308 --HGVQLKGGDQILLPQLSGDERKNACPMHVDFFSKVSHHTFGHSHLCLGQHL 365
Db 303 DLDGTVIPAGDVLVVLADAHRTPERFPDPHFRDTRDTAGHLAFGHGIFHCIGAP 362
QY 366 ELIIVLKELWLTIPDFS--IAPGAQIQKSGIVSGVQALPLVW 406
Db 363 EARIARALLERCPDLALDVSPGELVWYFNPIMRGLKALPIRW 405
```

RESULT 10

```
US-09-860-846-39
; Sequence 39, Application US/09860846
```

```
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/860,846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRP
; ORGANISM: Streptomyces venezuelae
US-09-860-846-39
```

```
Query Match 16.7%; Score 363.5; DB 9; Length 416;
Best Local Similarity 28.3%; Pred. No. 1.3e-27;
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;
```

```
QY 76 EDYRHSSECPPIPREAGEAYDFTSMDDPPQROFRALANQVGMVVDKLENRIQELA 135
Db 72 KDW--NSTPLTEAEALNHNWLES--DPPRHLRLKLVAREFTMRRVVELLRPRVQEI 127
QY 136 CSLIESL--RPOGQCNFTEDYAEPPPIEFMLLAGLPEEDIHKLXLTQMTDPPGSM 193
Db 128 DGLVDAMLAAPDGRADLMESLAWPLITWISLGVPEPDRAAFRWTDVAFVDDPAQA 187
QY 194 AEAKALYDYLPIIEQRKQKGTDAIS-IVANGQVNGRPIITSDEAKRMFGLLVGG 252
Db 188 QTAAEMSYLSRLIDSKRGQDGLLSALVTSDEGSRLLSEELGMAHILLVAGHET 247
QY 253 VVNF--SFMSEFLAKSPEHQELIERPERIPAAACEILLRFSIVADGRILTSYEF 307
Db 248 TYNLIANGMYALLSHPDQALRADMTLLDGAVEEMLR-----YEGPVESATYRFP 302
QY 308 --HGVQLKGGDQILLPQLSGDERKNACPMHVDFFSKVSHHTFGHSHLCLGQHL 365
Db 303 DLDGTVIPAGDVLVVLADAHRTPERFPDPHFRDTRDTAGHLAFGHGIFHCIGAP 362
QY 366 ELIIVLKELWLTIPDFS--IAPGAQIQKSGIVSGVQALPLVW 406
Db 363 EARIARALLERCPDLALDVSPGELVWYFNPIMRGLKALPIRW 405
```

RESULT 11

```
US-09-988-384B-39
; Sequence 39, Application US/0988384B
; Publication No. US2003007382A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.536US1
; CURRENT APPLICATION NUMBER: US/09/988,384B
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: PCT/US99/14398
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRP
; ORGANISM: Streptomyces venezuelae
US-09-988-384B-39
```


; Sequence 13, Application US/10201365
; Publication No. US20030148469A1
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: COMBINATORIAL POLYKETIDE LIBRARIES PRODUCED USING A MODULAR
; FILE OF INVENTION: PKS GENE CLUSTER AS SCAFFOLD
; FILE REFERENCE: 300622002103
; CURRENT APPLICATION NUMBER: US/10/201.365
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/141,908
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: US 09/073,538
; PRIOR FILING DATE: 1998-05-06
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 416
; TYPE: PR1
; ORGANISM: Streptomyces venezuelae
US-10-201-365-13

Query Match 15.7%; Score 363.5; DB 14; Length 416;
Best Local Similarity 28.3%; Pred. No. 1.3e-27;
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

QY 76 EDYRHSSECFIPREAGEAYDFIPTSMDPPEQRFALANQVVGMPVVDKLENRIQELA 135
Db 72 KDWK--NSTYPLTEAEALNHNMLE--DPRHTRLRKLVAAREFTMRVVELLRPRVQEIIV 127

QY 136 CSLTESL--RPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYLTDQWTRDGSMTFF 193
Db 128 DGLVDMLAAPDGRADLMESLAWPLPITVISELLGVPEPDRAAFRVWTDAFVFPDPAQA 187

QY 194 ABAKEALDYLIPIEORRQKPGTDAIS--IVANGCVNGRPITSDEAKEMFGILLVGGLD 252
Db 188 QTAWAEMSGYLSRLIDSKRGQDGLLSALVTSDEDSRLTSELLGNAHILLVAGHET 247

QY 253 VVNFLSFMEFLAKSPHRRQELIERPERIPAAACELLRRFSLVADGRILTSDFEYF----- 307
Db 248 TVNLIANGMYALLSHPDQALRADMTLLDGAVEMLR-----YEGPVESATYRFPVEPV 302

QY 308 --HGVQLKKGOIILLPQMLSGLDERKNACPMVEVDFSRKVSHTTTFGHSHLCLGQHARR 365
Db 303 DLDCGTVPAGDTVLVVLADAHRTPERFDPHRTDRTAGHLAFGHGIFHCIGAPLARR 362

QY 366 EIIIVTLKEMWTRIPDFS--IAPGAQIOHKSGIVSGVQALPLVW 406
Db 363 EARIATVALLERCPDLALDVSFGLVWY?NPMIRGLKALPIRW 405

RESULT 15

US-10-160-539-18
; Sequence 18, Application US/10160539
; Publication No. US20030162262A1
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE OF INVENTION: 300622002120
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/10/-60.539
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: US/09/657,440
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/320,878
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: CIP OF 09/141,908

; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 416
; TYPE: PR1
; ORGANISM: Streptomyces venezuelae
US-10-160-539-18

Query Match 16.7%; Score 363.5; DB 14; Length 416;
Best Local Similarity 28.3%; Pred. No. 1.3e-27;
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;

QY 76 EDYRHSSECFIPREAGEAYDFIPTSMDPPEQRFALANQVVGMPVVDKLENRIQELA 135
Db 72 KDWK--NSTYPLTEAEALNHNMLE--DPRHTRLRKLVAAREFTMRVVELLRPRVQEIIV 127

QY 136 CSLTESL--RPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYLTDQWTRDGSMTFF 193
Db 128 DGLVDMLAAPDGRADLMESLAWPLPITVISELLGVPEPDRAAFRVWTDAFVFPDPAQA 187

QY 194 ABAKEALDYLIPIEORRQKPGTDAIS--IVANGCVNGRPITSDEAKEMFGILLVGGLD 252
Db 188 QTAWAEMSGYLSRLIDSKRGQDGLLSALVTSDEDSRLTSELLGNAHILLVAGHET 247

QY 253 VVNFLSFMEFLAKSPHRRQELIERPERIPAAACELLRRFSLVADGRILTSDFEYF----- 307
Db 248 TVNLIANGMYALLSHPDQALRADMTLLDGAVEMLR-----YEGPVESATYRFPVEPV 302

QY 308 --HGVQLKKGOIILLPQMLSGLDERKNACPMVEVDFSRKVSHTTTFGHSHLCLGQHARR 365
Db 303 DLDCGTVPAGDTVLVVLADAHRTPERFDPHRTDRTAGHLAFGHGIFHCIGAPLARR 362

QY 366 EIIIVTLKEMWTRIPDFS--IAPGAQIOHKSGIVSGVQALPLVW 406
Db 363 EARIATVALLERCPDLALDVSFGLVWY?NPMIRGLKALPIRW 405

Search completed: April 6, 2004, 19:14:31
Job time : 35.974 secs

QY 253 VVNFSLFSMEFLAKSPHQRQLERPERIPAAACELLRRFSLVADGRILTSYEF----- 307
Db 248 TVNLIANGNYALLSHPDQALRADMTLLDGAVEMLR-----YEGPVESATYRPPVEPV 302
QY 308 --HGVLKKGQDQILLPQMLSGDERKNACPMHVDFSRQKVSHTTFHGHSHLCIGQHARR 365
Db 303 DLGDTVIPAGDTVLVLAADAHRTPERFDPHREDIRDUTAGHLAFGHGHCIGAPLARR 362
QY 366 EIIIVTLKEWLRTPDFS--IAPGAQIOHKSIGVSGVQALPLVW 406
Db 363 EARIAVALLERCPLDLDVSPGELVWYVNPMPMIRGLKALPIRW 405

RESULT 2
US-09-105-537-39
; Sequence 39, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xie, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105.537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version: 3.0
; SEQ ID NO 39
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-105-537-39

Query Match 16.7%; Score 363.5; DB 3; Length 416;
Best Local Similarity 28.3%; Pred. No. 5.5e-29;
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;
QY 76 EDYRHSSECFPIREAGEAYDFPTSDPPEQRFALANQVGVVVDKLNRIQELA 135
Db 72 KDMR--NSTTPTLTERAALNENMLES--DPPRHTLRKLKLVAREFTMRVRLPRVQEV 127
QY 136 CSLIESL--RPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYLTDQMTDPGSMTF 193
Db 128 DGLVDMLAAPDGRADIMESLAWPLPITVISSELLGVPEPDRAAERWTDVFPDDPAQA 187
QY 194 AEAEKALVDYLIPIEORRQKPGTDAIS--IVANGQVNGRPITSDRAKMFGLLLVGGD 252
Db 188 QTAMAENGSLRLSDSKRGQDGEDLLSALVRTSDEGSRLTSELLGMAHILLVAGHET 247
QY 253 VVNFSLFSMEFLAKSPHQRQLERPERIPAAACELLRRFSLVADGRILTSYEF----- 307
Db 248 TVNLIANGNYALLSHPDQALRADMTLLDGAVEMLR-----YEGPVESATYRPPVEPV 302
QY 308 --HGVLKKGQDQILLPQMLSGDERKNACPMHVDFSRQKVSHTTFHGHSHLCIGQHARR 365
Db 303 DLGDTVIPAGDTVLVLAADAHRTPERFDPHREDIRDUTAGHLAFGHGHCIGAPLARR 362
QY 366 EIIIVTLKEWLRTPDFS--IAPGAQIOHKSIGVSGVQALPLVW 406
Db 363 EARIAVALLERCPLDLDVSPGELVWYVNPMPMIRGLKALPIRW 405

RESULT 3
US-09-141-908-13
; Sequence 13, Application US/09141908
; Patent No. 6503741
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li

; TITLE OF INVENTION: Combinatorial Polyketide Libraries Produced Using a
; FILE OF INVENTION: Modular PKS Gene Cluster as Scaffold
; CURRENT APPLICATION NUMBER: US/09/141.908
; CURRENT FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073.538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846.247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: PROV. 60/376.919
; EARLIER FILING DATE: 1998-03-05
; EARLIER APPLICATION NUMBER: PROV. 60/387.080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-141-908-13

Query Match 16.7%; Score 363.5; DB 4; Length 416;
Best Local Similarity 28.3%; Pred. No. 5.5e-29;
Matches 97; Conservative 62; Mismatches 163; Indels 21; Gaps 7;
QY 76 EDYRHSSECFPIREAGEAYDFPTSDPPEQRFALANQVGVVVDKLNRIQELA 135
Db 72 KDMR--NSTTPTLTERAALNENMLES--DPPRHTLRKLKLVAREFTMRVRLPRVQEV 127
QY 136 CSLIESL--RPQGCNFTEDYAEPPPIRIFMLLAGLPEEDIPHLKYLTDQMTDPGSMTF 193
Db 128 DGLVDMLAAPDGRADIMESLAWPLPITVISSELLGVPEPDRAAERWTDVFPDDPAQA 187
QY 194 AEAEKALVDYLIPIEORRQKPGTDAIS--IVANGQVNGRPITSDRAKMFGLLLVGGD 252
Db 188 QTAMAENGSLRLSDSKRGQDGEDLLSALVRTSDEGSRLTSELLGMAHILLVAGHET 247
QY 253 VVNFSLFSMEFLAKSPHQRQLERPERIPAAACELLRRFSLVADGRILTSYEF----- 307
Db 248 TVNLIANGNYALLSHPDQALRADMTLLDGAVEMLR-----YEGPVESATYRPPVEPV 302
QY 308 --HGVLKKGQDQILLPQMLSGDERKNACPMHVDFSRQKVSHTTFHGHSHLCIGQHARR 365
Db 303 DLGDTVIPAGDTVLVLAADAHRTPERFDPHREDIRDUTAGHLAFGHGHCIGAPLARR 362
QY 366 EIIIVTLKEWLRTPDFS--IAPGAQIOHKSIGVSGVQALPLVW 406
Db 363 EARIAVALLERCPLDLDVSPGELVWYVNPMPMIRGLKALPIRW 405

RESULT 4
US-09-657-440-18
; Sequence 18, Application US/09657440
; Patent No. 6509455
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/657.440
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/320.878
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: CIP OF 09/141.908
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 416
; TYPE: PRT


```
TEPPERMAN, JAMES M.
; TITLE OF INVENTION: EXPRESSION OF HERBICIDE METABOLIZING
; CYTOCHROMES
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION NUMBER: US/07/569,781
; FILING DATE: 23-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 464,499
; FILING DATE: 12-JAN-1990
; APPLICATION NUMBER: 405,605
; FILING DATE: 11-SEP-1989
; SEQ ID NO.6:
; LENGTH: 406
5212296-6

Query Match 15.1%; Score 324; DB 6; Length 406;
Best Local Similarity 26.3%; Pred. No. 2.5e-25;
Matches 104; Conservative 74; Mismatches 193; Indels 24; Gaps 11;

QY 31 PSNLSAGVQ--EAMVQLQESNVPDLVWTRCNGGH-WIATRGQLIREAYEDYR----- 79
Db 17 ESNRSCPYQLPDGYAQLRDPGPJHRVTLYDGRQAWVTTKHEARKLLGDPRLSSNRIDD 76
QY 80 HFSECEFPFIP--REAGEAYDFIPTSDMPPEQORQFRALANQVGVNVDKLENRIQELACS 137
Db 77 NFATSPREFAVRESPOAF---IGLDPPEHGTERRMTISEFTWKRIKGRPEVEVVG 132
QY 138 LIESLRFQ--QCNFTEYAEFPFIRIPMLLAGPEEDIPHLKYLTDQMTDPGSMTFAEA 196
Db 133 FLDEMLAAGTADLVSQFALPVSMTICLLGVFPYADHEFFQASKRLVQSTDAQSALTA 192
QY 197 KEALYLYLPIIQRCKRGKGTDAI-SIVANGQNGRPITSDEAKMFGLLVGLDVTVN 255
Db 193 RNLGAYLDGLITQFQEPGAGLVGALVQDLANGE-IDREELSTAMLLLIAGHETTAS 251
QY 256 FLGFSMEFLAKSPEHQELTERPERIPAACEELRRFSL--VADGRILTSDFEFGVQLK 313
Db 252 MTSLSVITLLDHEQYAAALRADSLVPGAVEELLRYLAIAIADAGRVATADIEVEGLIR 311
QY 314 KGQQLPQWGLDERKNAACPMHVDPSRKVSHTTFGHSHLCLGQHLARREIIVTKE 373
Db 312 AGEGVIVNNGIANRGDVTVEDPOALDIHRSARHHLAFGFGVHQCGLQGLARLELEVI 371
QY 374 WLTRIPDFTA--GCAQTOHKSG-IVSGVQALPLVW 406
Db 372 LMDRVPTRLAVPVEQVLVRPGTTIQGVNELPVTM 406

RESULT 8
US-08-765-907A-10
; Sequence 10, Application US/08765907A
; Patent No. 6352839
; GENERAL INFORMATION:
; APPLICANT: BLANC, Veronique
; APPLICANT: THIBAUT, Denis
; APPLICANT: BAWAS-JACQUES, Nathalie
; APPLICANT: BLANCHE, Francis
; APPLICANT: COUZET, Joel
; APPLICANT: BARRIERE, Jean-Claude
; APPLICANT: DEBUSSCHE, Laurent
; APPLICANT: FAMECHON, Alain
; APPLICANT: PARIS, Jean-Marc
; APPLICANT: DUTRUC-ROSSET, Gilles
; TITLE OF INVENTION: Streptogramins And Method For Preparing Same By
; TITLE OF INVENTION: Mutasynthesis
; FILE REFERENCE: Streptogramin genes
; CURRENT APPLICATION NUMBER: US/08/765,907A
; CURRENT FILING DATE: 1997-03-20
; NUMBER OF SEQ IDS NOS: 17
; SOFTWARE: PatentID ver. 2.0
; SEQ ID NO 10
; LENGTH: 399

; TYPE: PRT
; ORGANISM: Streptomyces pristinaespiralis
US-08-765-907A-10

Query Match 14.9%; Score 324.5; DB 4; Length 399;
Best Local Similarity 29.0%; Pred. No. 5.7e-25;
Matches 106; Conservative 59; Mismatches 159; Indels 41; Gaps 12;

QY 74 AYEDYR-----FSSECFPIPREAGEAYDFIPTSDMPPEQORQFRALANQVGM 121
Db 36 AFHVERHADVLTVASDPGVYSSQLSRPQSALSEQLSVLDPFEMHRTLRLVSOQFTF 95
QY 122 PVVDKLENRIQELACSILSRPQGC--NETEDYAEFPFIRIPMLLAGPEEDIPHLKYL 180
Db 96 RTVADLEPRVTELQGLLDV--DGDTFDLVADFAVPLVIVIAELLGVPPADRTLFRSW 153
QY 181 TDQWTR-----PGSMTFAEAKALYDYLPIIPIEQRQRKGTDAISIVA 224
Db 154 SDRMLQWQVADPADMQFGDDADEYQRLVKPEPMEAHAYLHDHVTDRRAREPANDLSALV 213
QY 225 NGVNGCRPITSDEAKMFGLLVGLDVTVNFLSFSMEFLAKSPEHQELTERPER--IP 282
Db 214 AARVEGRLUDEQVEFGALLMAGHVSTSMLLGNVTCLXQHP--RAEAAARADRSIP 271
QY 283 AACBELR-RFSLVADGRILTSDFEFGVQ--KKGQDQILLPQMLS--GLDERKNAACPMHVD 340
Db 272 ALIEVLRLRPITVMAV--TKDTVLAGTTIPAG--RMVTPSLLSANHDEQVTFDPLDLD 330
QY 341 SRQKVSHTTGHSHLCLGQHLAREIIVLKEWLTRIPDPSIAPGAIQ--HKSGIVSGV 399
Db 331 AREG--RQIAFGHGHYCLGAPLARLEGRIALFALDFRPPDFSTGAKIRVHRDGLF--GV 388
QY 400 QALPL 404
Db 389 KNLPL 393

RESULT 9
US-09-385-028-12
; Sequence 12, Application US/09385028
; Patent No. 6232106
; GENERAL INFORMATION:
; APPLICANT: Susan E. Jensen
; APPLICANT: Kwamena A Aidoo
; APPLICANT: Ashish S. Paraskar
; TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic
; Patent No. 6232106
; TITLE OF INVENTION: Acid Biosynthesis
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN,PPLC
; STREET: The Jenifer Bulding, 400 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/385,028
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/790,462
; FILING DATE: 29-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: D. Douglas Price
; REGISTRATION NUMBER: 24,514
; REFERENCE/DOCKET NUMBER: 14-8/P57452US2
; TELECOMMUNICATION INFORMATION:
```

TELEPHONE: (202) 638-6666
TELEFAX: (202) 39305350
TELEX: RCA 248593 IDEA UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 409 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-385-028-12

Query Match 14.8%; Score 322; DB 3; Length 409;
Best Local Similarity 28.7%; Pred. No. 1.1e-24;
Matches 121; Conservative 60; Mismatches 186; Indels 54; Gaps 18;

QY 16 PHVEHLVDFDMMYNSNLGAGVCEAWAVLOESNVPLVWTRCNCGGH-WIATRGQLIREA 74
DB 14 PAYEMGRVCPVD---PPQLAGLSQKAAASRT-----LW---DGSQVWLVTSHAGARAV 62

QY 75 YEDYRHFS-SECPRIP-----REAGEAYDFIPTSMDDPPQCRPRA-----LANOV 118
DB 63 LGDRRTAVTSAPGFPMLTRISQLVANPESASFI--RMDDPQHSRLRSLMTRDFLARRA 120

QY 119 VGM-PVVDKLENRIQELACSLIESLRPGQCNFTEDYAEPPPIRIFMLLAGLPEDIPHL 177
DB 121 EALRPVAVREL---LDEILGGLVKGERP---VDLVAGLTTPVPSRVITLLFGAGDDRREFI 174

QY 178 K-----YLTQMTDPGSMTPFAEAKALVDYLIPIIEQRQKPGTDAISIVANGQVNGRPI 233
DB 175 EDRSAVLIDRGVYTP---QVAKARDELGDYLRLEIVEERIENPGTDLISRLVIDQVRPGHL 231

QY 234 TSDAEKMFGLLVGGLTVDVNF--SFSMEFLAKSPHEHQELIEREPIPAACEELLRRFS 293
DB 232 RVEEMVPMCRLLLVAGHGTTTSQASLSLLTDPDELAGLTEDPALLPKAVEELLRPHS 291

QY 294 LVADG--RLTSDYEFHGVOLKKGQDQILLPQMLSGLDERKNACPMHVDFFSRQKVSHTTFFG 351
DB 292 IVQNGLARAAVEDVQDDVLIRAGEGVVLSLSAGNRDETVPDPDRVDVDRDARRHLAFG 351

QY 352 HGHSLCGLQHLAR---REIIVTLKEWLTRIPDSIA-PGAQIOHKSIGVS-GVQALPLVM 406
DB 352 HGMHQCLGQWLARVELEILLAAVLRW---PGARLAVFPEELDFRHEVSSVGLGALPVTM 408

QY 407 D 407
DB 409 Z 409

RESULT 10
US-09-726-614-12
Sequence 12, Application US/09726614
Patent No. 6514735
GENERAL INFORMATION:
APPLICANT: Susan E. Jensen
APPLICANT: Kwamena A. Aidoo
APPLICANT: Ashish S. Paradkar
TITLE OF INVENTION: DNA Sequence Encoding Enzymes of Clavulanic
Patent No. 6514735
TITLE OF INVENTION: Acid Biosynthesis
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC
STREET: The Jennifer Building, 400 Seventh Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/726,614
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/790,462
FILING DATE: 29-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: D. Douglas Price
REGISTRATION NUMBER: 24,514
REFERENCE/DOCKET NUMBER: 1418/P57452US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 39305350
TELEX: RCA 248593 IDEA UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 409 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-726-614-12

Query Match 14.8%; Score 322; DB 4; Length 409;
Best Local Similarity 28.7%; Pred. No. 1.1e-24;
Matches 121; Conservative 60; Mismatches 186; Indels 54; Gaps 18;

QY 16 PHVEHLVDFDMMYNSNLGAGVCEAWAVLOESNVPLVWTRCNCGGH-WIATRGQLIREA 74
DB 14 PAYEMGRVCPVD---PPQLAGLSQKAAASRT-----LW---DGSQVWLVTSHAGARAV 62

QY 75 YEDYRHFS-SECPRIP-----REAGEAYDFIPTSMDDPPQCRPRA-----LANOV 118
DB 63 LGDRRTAVTSAPGFPMLTRISQLVANPESASFI--RMDDPQHSRLRSLMTRDFLARRA 120

QY 119 VGM-PVVDKLENRIQELACSLIESLRPGQCNFTEDYAEPPPIRIFMLLAGLPEDIPHL 177
DB 121 EALRPVAVREL---LDEILGGLVKGERP---VDLVAGLTTPVPSRVITLLFGAGDDRREFI 174

QY 178 K-----YLTQMTDPGSMTPFAEAKALVDYLIPIIEQRQKPGTDAISIVANGQVNGRPI 233
DB 175 EDRSAVLIDRGVYTP---QVAKARDELGDYLRLEIVEERIENPGTDLISRLVIDQVRPGHL 231

QY 234 TSDAEKMFGLLVGGLTVDVNF--SFSMEFLAKSPHEHQELIEREPIPAACEELLRRFS 293
DB 232 RVEEMVPMCRLLLVAGHGTTTSQASLSLLTDPDELAGLTEDPALLPKAVEELLRPHS 291

QY 294 LVADG--RLTSDYEFHGVOLKKGQDQILLPQMLSGLDERKNACPMHVDFFSRQKVSHTTFFG 351
DB 292 IVQNGLARAAVEDVQDDVLIRAGEGVVLSLSAGNRDETVPDPDRVDVDRDARRHLAFG 351

QY 352 HGHSLCGLQHLAR---REIIVTLKEWLTRIPDSIA-PGAQIOHKSIGVS-GVQALPLVM 406
DB 352 HGMHQCLGQWLARVELEILLAAVLRW---PGARLAVFPEELDFRHEVSSVGLGALPVTM 408

QY 407 D 407
DB 409 Z 409

RESULT 11
US-09-385-040-12
Sequence 12, Application US/09385040
Patent No. 6589775
GENERAL INFORMATION:
APPLICANT: Jensen, Susan E
APPLICANT: Aidoo, Kwamena A
APPLICANT: Paradkar, Ashish S
TITLE OF INVENTION: DNA SEQUENCE ENCODING ENZYMES OF CLAVULANIC ACID
TITLE OF INVENTION: BIOSYNTHESIS
FILE REFERENCE: 09/385,040
CURRENT APPLICATION NUMBER: US/09/385,040

```

; CURRENT FILING DATE: 1999-08-30
; PRIOR APPLICATION NUMBER: US 08/790,452
; PRIOR FILING DATE: 1997-01-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Streptomyces clavuligerus
US-09-385-040-12

Query Match      14.8%; Score 322; DB 4; Length 409;
Best Local Similarity 28.7%; Pred. No. 1.1e-24;
Matches 121; Conservative 60; Mismatches 186; Indels 54; Gaps 18;

QY 16 PHVHELVDFDMYFNSAGVQEAQAVLQESWPCLVWTRNGGH-WIATGQILIREA 74
DB 14 PAYEMHVCPEVD---PPQLAGLRSQKAASRVT---LW---DGSQVWLVTSAGARAV 62
QY 75 YEDYRHFS-SECFPE-----REAGEAYDFI27SMDDPEQROFRA-----LANQV 118
DB 63 LGDRRTAVTSAPGPFMLTSTLSOLVRANPESASFI--RMDDPQHSRLRSLMLTRDFLAR 120
QY 119 VGM-PVVDKLENRIOLACSLIESLRPQGCNFTEDYAEPPRIIFMLLAGLPEEDIPHL 177
DB 121 EALRPVAVREL---LDELGLVKGER---VDLVAGLITVPVSRVITLLEGAGDDRREFI 174
QY 178 K-----YLDQMTPEPGSGMTFAEAKALYV---PIIEQRRQKPGTDAISIVANGQVNGRPI 233
DB 175 EDRSAVLDIGRYTPE---QVAKARDELGYRLVBEERIEENPGTDLISRJVIDQVRPGHL 231
QY 234 TSDEAKEMFGLIAGVGLDVTWNFLSFSMEFLAKSPHQELIERPERIPAAACELLRRFS 293
DB 232 RVEEMVPMCLLVAGHGTT---SGASLSLSLLTDPELAGHLETDPAALLPAAVELLRFHS 291
QY 294 LVADG--RLTSDYEFHGVQKKGQDQILLPQMLSGLDERKXNACPMRVDFSRQKVSHTTFG 351
DB 292 IVQNGLARAAVEDVQDDVLIRAGEGVLSLSAGNRDETVEFPDVRVDVDRDARRHLAFG 351
QY 352 HSHLCIGQHAR---REIIVTKEMLTRIPDFSTIA-PGAQ-CHKSGIVS-GVOALPLVW 406
DB 352 HGMHQCGLGOWLARVELEELAAVLRW---PGARLAVFFELDFRHEVSSYGLGALPVTW 408
QY 407 D 407
DB 409 Z 409

RESULT 12
US-08-102-863-11
; Sequence 11, Application US/08102863
; Patent No. 5466590
; GENERAL INFORMATION:
; APPLICANT: SARIASLANI, SIMA
; TITLE OF INVENTION: CONSTITUTIVE
; TITLE OF INVENTION: EXPRESSION OF P450SOY
; TITLE OF INVENTION: AND FERREDOXIN-SOY IN
; TITLE OF INVENTION: STREPTOMYCES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: USA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0,
; SOFTWARE: Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/102,863
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/807,001
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: GALLEGO, R. THOMAS
; REGISTRATION NUMBER: 32,692
; REFERENCE/DOCKET NUMBER: CR-9000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-7342
; TELEFAX: 302-892-7949
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 412 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-102-863-11

Query Match      14.4%; Score 314; DB 1; Length 412;
Best Local Similarity 27.2%; Pred. No. 7.3e-24;
Matches 84; Conservative 56; Mismatches 163; Indels 6; Gaps 5;

QY 103 MDPEQRFALANOVGMVVDKLENRIOLACSLIESLRPOG-QCNFTEDYAEPPIR 161
DB 105 VDEPHNTQRRMLIPTFSVKRIGALREIQTVDRLDAMERQGPFAELVSFAFALPVFSM 164
QY 162 IFMLLAGLPEEDIPHLKYLTDQMTPEPGSGMTFAEAKALYDYLPIIEQRRQKPGTDAIS 221
DB 165 VICALLGVVADHAFPEERSQRLLEGAGDVNARDELEBYLGALIDRKAEPGCGLLD 224
QY 222 IVANGQVNGRPTSDENKRMFGLLVGLDVTWNFLSFSMEFLAKSPHQELIERPERI 281
DB 225 ELIHRDHPDGPVDRQELVAFVILLIAGHETTANMISLGTFTLLSHPEQLAALRAGTST 284
QY 282 PAACEELLRFSLVADG--RLTSDYEFHGVQKKGQDQILLPQMLSGLDERKXNACPMHVD 339
DB 285 AVVVEELL-RFLSTAEGLQRLATEDMEVDGATIRKGGVWFVSTSLINRDADVFPRAETLD 343
QY 340 FSRQKVSHTTFHGHSHLCIGQHARREIIVTKEMLTRIPDFSTIA-PCAQIOHKSG-IVS 397
DB 344 WDRPARHHLAFGFGVHQICQNLARAELELDIAMTLFERLPGLR-LAVPAHEIRHKPGDTIQ 403
QY 398 GVOALPLVW 406
DB 404 GLDLPVAV 412

RESULT 13
PCT-US92-10885-11
; Sequence 11, Application PC/TUS9210885
; GENERAL INFORMATION:
; APPLICANT: SARIASLANI, SIMA
; TITLE OF INVENTION: CONSTITUTIVE
; TITLE OF INVENTION: EXPRESSION OF P450SOY
; TITLE OF INVENTION: AND FERREDOXIN-SOY IN
; TITLE OF INVENTION: STREPTOMYCES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: USA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch,
; MEDIUM TYPE: 1.0 MB
```


Wed Apr 7 09:15:31 2004

QY 6 QSNANLAPLPHVPEHLVDFDXNPSNLSAGVQE-----AWAVLQSNVDPDLVWTRCNGG 61
Db 3 QEQANQSEIKP-----AEDFKFPAP-----GYAEDPFPAPAESLREA-TFIFYWD--EGR 48
QY 62 HWIATRQ-----LREAYDYHFSSECEPFIPIREACEAYDEIPTSMDDPEQR 103
Db 49 SWLTRYHVSFAVFRDERFAVSEESSAYSAIP-----ELSDMKYGLFGLPPEDHA 104
QY 110 QFRALANQVGMPPVVDKLENRIQELACSLIESLRPOQCNCFTEDYABPFPFIRIFMLLAGL 169
Db 105 RVKLVNPSFTSRAIDLRLRAEIQTVDQLDARSQSEEDVWRDYAESIPMRAISALKV 164
QY 170 PEEDIPHLKYLTDQMTDPDGSMTFAKAKALYKLP-----I 207
Db 165 PAE-----CDEKFRFGSAT-----ARALGVGLVPOVDEETKLVASVTEGLALLHDV 212
QY 208 IEORRQXP-GTDAISIVANGQNGRPITSDAKRMFGELLVGGLDTVVNFLSFSMEELAK 266
Db 213 LDERRRNPLENDVLTMLQAADGSRUSKELVALVGAITTAAGTDTTTLIAFAVNLJR 272
QY 267 SPEHRQELIERIPAAACEELLRRFSLVADG--RILTSDFEHGVOLKKGDOI--LLPQ 322
Db 273 SPEALELYKABEGLMRNALDEVRFDNLRIGTVFARQDLEYCGASIKKGMVELLIPS 332
QY 323 MLSGLDERKNACPMHVDPSRQKVSHTTFGHSHLCGQHLARREIIVTLKEWLTIPDPS 382
Db 333 ALR--DGTVFSRPOVDFVRRD*GASLAYGRGPHVCPGVSLARLEABIAVGTIFRRFPENK 390
QY 383 - 383
Db 391 - 391

Search completed: April 6, 2004, 18:56:00
Job time : 15.9588 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: April 6, 2004, 18:53:57 ; Search time 26.1039 Seconds
(without alignments)
3108.883 Million cell updates/sec

Title: US-09-246-451A-17

Perfect score: 1508

Sequence: 1 MQLTFFYNVNSCPVNSNIVR.....PLTGTGGQIRLNCRVNSNS 309

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 26263353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/ptodata1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1608	100.0	309	15 US-10-453-104-17	Sequence 17, Appl
2	1603	99.7	309	14 US-10-274-793-4	Sequence 4, Appl
3	1601	99.6	309	14 US-10-274-793-6	Sequence 6, Appl
4	1586	99.3	309	14 US-10-274-793-18	Sequence 18, Appl
5	1595	99.2	309	14 US-10-274-793-8	Sequence 8, Appl
6	1595	99.2	309	14 US-10-274-793-10	Sequence 10, Appl
7	1594	99.1	309	14 US-10-274-793-12	Sequence 12, Appl
8	1590	98.9	309	14 US-10-274-793-14	Sequence 14, Appl
9	1577	98.1	309	14 US-10-274-793-20	Sequence 20, Appl
10	1577	98.1	309	14 US-10-274-793-24	Sequence 24, Appl
11	1572	97.8	309	14 US-10-274-793-22	Sequence 22, Appl
12	921	57.3	352	12 US-10-424-599-278555	Sequence 278555,
13	912.5	56.7	354	12 US-10-424-599-250229	Sequence 250229,
14	885	55.0	353	12 US-10-424-599-251628	Sequence 251628,
15	874.5	54.4	331	12 US-10-424-599-276196	Sequence 276196,

16	860.5	53.5	333	14 US-10-259-165-18	Sequence 18, Appl
17	860.5	53.5	333	14 US-10-259-165-370	Sequence 370, Appl
18	846.5	52.6	352	12 US-10-424-599-249564	Sequence 249564,
19	836.5	52.0	349	12 US-10-425-114-58420	Sequence 58420, A
20	813	50.6	358	12 US-10-424-599-159302	Sequence 159302,
21	799	49.7	339	12 US-10-425-114-47322	Sequence 47322, A
22	781	48.6	344	14 US-10-289-757-111	Sequence 111, Appl
23	779	48.4	330	12 US-10-424-599-199360	Sequence 199360,
24	770.5	47.9	334	14 US-10-259-165-168	Sequence 168, Appl
25	767.5	47.7	335	16 US-10-389-566-716	Sequence 716, Appl
26	767.5	47.7	335	16 US-10-389-566-2314	Sequence 2314, Ap
27	767.5	47.7	339	14 US-10-259-165-60	Sequence 60, Appl
28	767.5	47.7	339	14 US-10-259-165-396	Sequence 396, Appl
29	767.5	47.7	333	16 US-10-389-566-549	Sequence 549, Appl
30	765.5	47.6	329	16 US-10-389-566-1905	Sequence 1905, Ap
31	758.5	47.2	338	16 US-10-389-566-1496	Sequence 1496, Ap
32	758.5	47.2	338	16 US-10-389-566-1602	Sequence 1602, Ap
33	757.5	47.1	333	16 US-10-389-566-891	Sequence 891, Appl
34	755.5	47.0	334	16 US-10-389-566-2367	Sequence 2367, Ap
35	744.5	46.3	333	16 US-10-389-566-869	Sequence 869, Appl
36	741.5	46.1	316	12 US-10-424-599-279090	Sequence 279090,
37	737.5	45.9	339	16 US-10-389-566-658	Sequence 658, Appl
38	735	45.7	347	12 US-10-424-599-247654	Sequence 247654,
39	733.5	45.6	337	16 US-10-389-566-1239	Sequence 1239, Ap
40	733.5	45.6	337	16 US-10-389-566-2366	Sequence 2366, Ap
41	733	45.6	276	12 US-10-425-114-47525	Sequence 47525, A
42	730.5	45.4	329	16 US-10-389-566-2383	Sequence 2383, Ap
43	728.5	45.3	320	12 US-10-424-599-184544	Sequence 184544,
44	727.5	45.2	315	12 US-10-424-599-246059	Sequence 246059,
45	727.5	45.2	318	14 US-10-174-693-389	Sequence 389, Appl

ALIGNMENTS

RESULT 1
US-10-453-104-17
; Sequence 17, Application US/10453104
; Publication No. US20030207345A1
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology;
; APPLICANT: Frances H. Arnold
; APPLICANT: Hyun Joo
; TITLE OF INVENTION: Oxygenase Enzymes and Screening Method
; FILE REFERENCE: 4358/1B827-US3
; CURRENT APPLICATION NUMBER: US/10/453,104
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 09/661,093
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 09/246,451
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: US 60/086,206
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/106,834
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-453-104-17

Query Match 100.0%; Score 1608; DB 15; Length 309;
Best Local Similarity 100.0%; Pred. No. 7.5e-163;
Matches 309; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQLTFFYNVNSCPVNSNIVRDTVNEIUSDPRIAASTRLHFHDCFVNGCDASILLNNTT 60
|||||

Db 1 MQLTPTFYDSCNVSNIVREDTIVNELRSDPRIAASILRLHFHDFCVNGCDASILLDNNTT 60
Qy 61 SFTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLITIAAQQSVTLAGGPSWRV 120
Db 61 SFTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLITIAAQQSVTLAGGPSWRV 120
Qy 121 PLGRDSLOAFDLANANLPAFPFTLPOLKDSFNVGLNRSSDLVALSGGHTFGKNQCRF 180
Db 121 PLGRDSLOAFDLANANLPAFPFTLPOLKDSFNVGLNRSSDLVALSGGHTFGKNQCRF 180
Qy 181 IMBRLYNFSNTGLPDPTLNTTYLQTLRGLCPNLGNLSALVDFDLRTPTIFDNKYVNL 240
Db 181 IMBRLYNFSNTGLPDPTLNTTYLQTLRGLCPNLGNLSALVDFDLRTPTIFDNKYVNL 240
Qy 241 OKGLIQSDQELFSSPDATDIPLVRSFANSTQTFNFVAFVEMDRMGNTPLTGTGGQIRL 300
Db 241 OKGLIQSDQELFSSPDATDIPLVRSFANSTQTFNFVAFVEMDRMGNTPLTGTGGQIRL 300
Qy 301 NCRVNSNS 309
Db 301 NCRVNSNS 309
RESULT 2
US-10-274-793-4
; Sequence 4, Application US/10274793
; Publication No. US20030153042A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY;
; APPLICANT: Frances S. ARNOLD
; APPLICANT: Zhanglin Lin
; TITLE OF INVENTION: Expression of Functional Eukaryotic
; FILE REFERENCE: 3369/1E804-US3
; CURRENT APPLICATION NUMBER: US/10/274,793
; PRIOR FILING DATE: 2002-10-21
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/538,591
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/247,232
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-274-793-4

Query Match 99.7%; Score 1603; DB 14; Length 309;
Best Local Similarity 99.7%; Pred. No. 2.6e-162;
Matches 309; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MQLTPTFYDSCNVSNIVREDTIVNELRSDPRIAASILRLHFHDFCVNGCDASILLDNNTT 60
Db 1 MQLTPTFYDSCNVSNIVREDTIVNELRSDPRIAASILRLHFHDFCVNGCDASILLDNNTT 60
Qy 61 SFTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLITIAAQQSVTLAGGPSWRV 120
Db 61 SFTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLITIAAQQSVTLAGGPSWRV 120
Qy 121 PLGRDSLOAFDLANANLPAFPFTLPOLKDSFNVGLNRSSDLVALSGGHTFGKNQCRF 180
Db 121 PLGRDSLOAFDLANANLPAFPFTLPOLKDSFNVGLNRSSDLVALSGGHTFGKNQCRF 180
Qy 181 IMBRLYNFSNTGLPDPTLNTTYLQTLRGLCPNLGNLSALVDFDLRTPTIFDNKYVNL 240
Db 181 IMBRLYNFSNTGLPDPTLNTTYLQTLRGLCPNLGNLSALVDFDLRTPTIFDNKYVNL 240

Qy 241 OKGLIQSDQELFSSPDATDIPLVRSFANSTQTFNFVAFVEMDRMGNTPLTGTGGQIRL 300
Db 241 OKGLIQSDQELFSSPDATDIPLVRSFANSTQTFNFVAFVEMDRMGNTPLTGTGGQIRL 300
Qy 301 NCRVNSNS 309
Db 301 NCRVNSNS 309
RESULT 3
US-10-274-793-6
; Sequence 6, Application US/10274793
; Publication No. US20030153042A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY;
; APPLICANT: Frances S. ARNOLD
; APPLICANT: Zhanglin Lin
; TITLE OF INVENTION: Expression of Functional Eukaryotic
; FILE REFERENCE: 3369/1E804-US3
; CURRENT APPLICATION NUMBER: US/10/274,793
; PRIOR FILING DATE: 2002-10-21
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/538,591
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/247,232
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-274-793-6

Query Match 99.6%; Score 1601; DB 14; Length 309;
Best Local Similarity 99.4%; Pred. No. 4.2e-162;
Matches 307; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MQLTPTFYDSCNVSNIVREDTIVNELRSDPRIAASILRLHFHDFCVNGCDASILLDNNTT 60
Db 1 MQLTPTFYDSCNVSNIVREDTIVNELRSDPRIAASILRLHFHDFCVNGCDASILLDNNTT 60
Qy 61 SFTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLITIAAQQSVTLAGGPSWRV 120
Db 61 SFTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLITIAAQQSVTLAGGPSWRV 120
Qy 121 PLGRDSLOAFDLANANLPAFPFTLPOLKDSFNVGLNRSSDLVALSGGHTFGKNQCRF 180
Db 121 PLGRDSLOAFDLANANLPAFPFTLPOLKDSFNVGLNRSSDLVALSGGHTFGKNQCRF 180
Qy 181 IMBRLYNFSNTGLPDPTLNTTYLQTLRGLCPNLGNLSALVDFDLRTPTIFDNKYVNL 240
Db 181 IMBRLYNFSNTGLPDPTLNTTYLQTLRGLCPNLGNLSALVDFDLRTPTIFDNKYVNL 240
Qy 241 OKGLIQSDQELFSSPDATDIPLVRSFANSTQTFNFVAFVEMDRMGNTPLTGTGGQIRL 300
Db 241 OKGLIQSDQELFSSPDATDIPLVRSFANSTQTFNFVAFVEMDRMGNTPLTGTGGQIRL 300
Qy 301 NCRVNSNS 309
Db 301 NCRVNSNS 309
RESULT 4
US-10-274-793-18
; Sequence 18, Application US/10274793

```
; Publication No. US20030153042A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY;
; APPLICANT: Frances S. ARNOLD
; APPLICANT: Zhonglin LIN
; TITLE OF INVENTION: Expression of Functional Eukaryotic
; TITLE OF INVENTION: proteins
; FILE REFERENCE: 3369/1E804-US3
; CURRENT APPLICATION NUMBER: US/10/274,793
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/654,493
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/538,591
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/247,232
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-274-793-18

Query Match          99.3%; Score 1596; DB 14; Length 309;
Best Local Similarity 99.4%; Pred. No. 1.4e-161;
Matches 307; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MQLTPTFDNSCPNVNIVRDTIYNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTT 60
Db 1 MQLTPTFDNSCPNVNIVRDTIYNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTT 60
Qy 61 SFTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTAAQOSVTLAGGPSNRV 120
Db 61 SFTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTAAQOSVTLAGGPSNRV 120
Qy 121 PLGRDLSQAFPLDLANANLPAPFTTLPLQKDSFRNVGLNRSSDLVALSGGHTFGKNQCRF 180
Db 121 PLGRDLSQAFPLDLANANLPAPFTTLPLQKDSFRNVGLNRSSDLVALSGGHTFGKNQCRF 180
Qy 181 IMDRLYNFSNTGLDPDPTINTTYLQTLRGLCPNGNLSALVDFDLRTPTIFDKKYVNLEE 240
Db 181 IMDRLYNFSNTGLDPDPTINTTYLQTLRGLCPNGNLSALVDFDLRTPTIFDKKYVNLEE 240
Qy 241 QKGLIQSDQELFSSPDATDTIPLVRSFANSQTTFNFAFVEMDRMGNTITPLTGTGGQIRL 300
Db 241 QKGLIQSDQELFSSPDATDTIPLVRSFANSQTTFNFAFVEMDRMGNTITPLTGTGGQIRL 300
Qy 301 NCRVNSNS 309
Db 301 NCRVNSNS 309

RESULT 5
US-10-274-793-8
; Sequence 8, Application US/10274793
; Publication No. US20030153042A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY;
; APPLICANT: Frances S. ARNOLD
; APPLICANT: Zhonglin LIN
; TITLE OF INVENTION: Expression of Functional Eukaryotic
; TITLE OF INVENTION: proteins
; FILE REFERENCE: 3369/1E804-US3
; CURRENT APPLICATION NUMBER: US/10/274,793
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/654,493
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/538,591
```

```
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/247,232
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-274-793-8

Query Match          99.2%; Score 1595; DB 14; Length 309;
Best Local Similarity 99.0%; Pred. No. 1.8e-161;
Matches 306; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MQLTPTFDNSCPNVNIVRDTIYNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTT 60
Db 1 MQLTPTFDNSCPNVNIVRDTIYNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTT 60
Qy 61 SFTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTAAQOSVTLAGGPSNRV 120
Db 61 SFTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTAAQOSVTLAGGPSNRV 120
Qy 121 PLGRDLSQAFPLDLANANLPAPFTTLPLQKDSFRNVGLNRSSDLVALSGGHTFGKNQCRF 180
Db 121 PLGRDLSQAFPLDLANANLPAPFTTLPLQKDSFRNVGLNRSSDLVALSGGHTFGKNQCRF 180
Qy 181 IMDRLYNFSNTGLDPDPTINTTYLQTLRGLCPNGNLSALVDFDLRTPTIFDKKYVNLEE 240
Db 181 IMDRLYNFSNTGLDPDPTINTTYLQTLRGLCPNGNLSALVDFDLRTPTIFDKKYVNLEE 240
Qy 241 QKGLIQSDQELFSSPDATDTIPLVRSFANSQTTFNFAFVEMDRMGNTITPLTGTGGQIRL 300
Db 241 QKGLIQSDQELFSSPDATDTIPLVRSFANSQTTFNFAFVEMDRMGNTITPLTGTGGQIRL 300
Qy 301 NCRVNSNS 309
Db 301 NCRVNSNS 309

RESULT 6
US-10-274-793-13
; Sequence 10, Application US/10274793
; Publication No. US20030153042A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY;
; APPLICANT: Frances S. ARNOLD
; APPLICANT: Zhonglin LIN
; TITLE OF INVENTION: Expression of Functional Eukaryotic
; TITLE OF INVENTION: proteins
; FILE REFERENCE: 3369/1E804-US3
; CURRENT APPLICATION NUMBER: US/10/274,793
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/654,493
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/538,591
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/247,232
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Escherichia coli
```

US-10-274-793-10

Query Match 99.2%; Score 1595; DB 14; Length 309;
Best Local Similarity 99.0%; Pred. No. 1.8e-161;
Matches 306; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MOLTTFYDNCSPVNSIVRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTT 60
DB 1 MOLTTFYDNCSPVNSIVRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTT 60
QY 61 SPTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGFSWRV 120
DB 61 SPTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGFSWRV 120
QY 121 PLGRDLSLOAFDLANANLPAPFFTLPOLKDSFRNVGLNRSSDLVALSGHTFGKNQCRF 180
DB 121 PLGRDLSLOAFDLANANLPAPFFTLPOLKDSFRNVGLNRSSDLVALSGHTFGKNQCRF 180
QY 181 IMDRLYNFSNTGLPDPFTLNTTYLQTLRGLCPNGNLSALVDFDLRTPTTFDNKYYVNL 240
DB 181 IMDRLYNFSNTGLPDPFTLNTTYLQTLRGLCPNGNLSALVDFDLRTPTTFDNKYYVNL 240
QY 241 QKGLIQSDOELEFSSPNATDTIPLVRSFANSTQTFNFAFVEMDRMGNTITPLTGTQGI 300
DB 241 QKGLIQSDOELEFSSPNATDTIPLVRSFANSTQTFNFAFVEMDRMGNTITPLTGTQGI 300
QY 301 NCRVNSNS 309
DB 301 NCRVNSNS 309

RESULT 7

US-10-274-793-12
; Sequence 12, Application US/10274793
; Publication No. US20030153042A1

; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY;
; APPLICANT: Frances S. ARNOLD
; APPLICANT: Zhanglin Lin
; TITLE OF INVENTION: Expression of Functional Eukaryotic
; FILE REFERENCE: 3369/18834-US3
; CURRENT APPLICATION NUMBER: US/10/274,793
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/654,493
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/538,591
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/247,232
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12

; LENGTH: 309

; TYPE: PRT

; ORGANISM: Escherichia coli

US-10-274-793-12

Query Match 99.1%; Score 1594; DB 14; Length 309;
Best Local Similarity 99.0%; Pred. No. 2.4e-161;
Matches 306; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MOLTTFYDNCSPVNSIVRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTT 60
DB 1 MOLTTFYDNCSPVNSIVRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTT 60
QY 61 SPTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGFSWRV 120
DB 61 SPTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGFSWRV 120

QY 121 PLGRDLSLOAFDLANANLPAPFFTLPOLKDSFRNVGLNRSSDLVALSGHTFGKNQCRF 180
DB 121 PLGRDLSLOAFDLANANLPAPFFTLPOLKDSFRNVGLNRSSDLVALSGHTFGKNQCRF 180
QY 181 IMDRLYNFSNTGLPDPFTLNTTYLQTLRGLCPNGNLSALVDFDLRTPTTFDNKYYVNL 240
DB 181 IMDRLYNFSNTGLPDPFTLNTTYLQTLRGLCPNGNLSALVDFDLRTPTTFDNKYYVNL 240
QY 241 QKGLIQSDOELEFSSPNATDTIPLVRSFANSTQTFNFAFVEMDRMGNTITPLTGTQGI 300
DB 241 QKGLIQSDOELEFSSPNATDTIPLVRSFANSTQTFNFAFVEMDRMGNTITPLTGTQGI 300
QY 301 NCRVNSNS 309
DB 301 NCRVNSNS 309

RESULT 8

US-10-274-793-14
; Sequence 14, Application US/10274793
; Publication No. US20030153042A1

; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY;
; APPLICANT: Frances S. ARNOLD
; APPLICANT: Zhanglin Lin
; TITLE OF INVENTION: Expression of Functional Eukaryotic
; FILE REFERENCE: 3369/18834-US3
; CURRENT APPLICATION NUMBER: US/10/274,793
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/654,493
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/538,591
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/247,232
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-274-793-14

- Query Match 98.9%; Score 1590; DB 14; Length 309;
Best Local Similarity 99.0%; Pred. No. 6.3e-161;
Matches 306; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MOLTTFYDNCSPVNSIVRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTT 60
DB 1 MOLTTFYDNCSPVNSIVRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLDNTT 60
QY 61 SPTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGFSWRV 120
DB 61 SPTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGFSWRV 120
QY 121 PLGRDLSLOAFDLANANLPAPFFTLPOLKDSFRNVGLNRSSDLVALSGHTFGKNQCRF 180
DB 121 PLGRDLSLOAFDLANANLPAPFFTLPOLKDSFRNVGLNRSSDLVALSGHTFGKNQCRF 180
QY 181 IMDRLYNFSNTGLPDPFTLNTTYLQTLRGLCPNGNLSALVDFDLRTPTTFDNKYYVNL 240
DB 181 IMDRLYNFSNTGLPDPFTLNTTYLQTLRGLCPNGNLSALVDFDLRTPTTFDNKYYVNL 240
QY 241 QKGLIQSDOELEFSSPNATDTIPLVRSFANSTQTFNFAFVEMDRMGNTITPLTGTQGI 300
DB 241 QKGLIQSDOELEFSSPNATDTIPLVRSFANSTQTFNFAFVEMDRMGNTITPLTGTQGI 300

```
QY 301 NCRVNSNS 309
Db 301 NCRVNSNS 309

RESULT 9
US-10-274-793-20
; Sequence 20, Application US/10274793
; Publication No. US20030153042A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY;
; APPLICANT: Frances S. ARNOLD
; APPLICANT: Zhanglin LIN
; TITLE OF INVENTION: Expression of Functional Eukaryotic
; TITLE OF INVENTION: proteins
; FILE REFERENCE: 3369/1E804-US3
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/654,493
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/538,591
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/247,232
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-274-793-20

Query Match 98.1%; Score 1577; DB 14; Length 309;
Best Local Similarity 98.4%; Pred. No. 1.5e-159;
Matches 304; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MQLPTFYDNCSPNVNIVRDTIVNELRSPRIAASILRLHFHDFCVNGCDASILLDNTT 60
Db 1 MQLPTFYDNCSPNVNIVRDTIVNELRSPRIAASILRLHFHDFCVNGCDASILLDNTT 60

QY 61 SFRTEKDAFGNANSARGFPVIDRMKAAVESACPRVTSCADLLIAAQQSVTLAGGFSWRV 120
Db 61 SFRTEKDAFGNANSARGFPVIDRMKAAVESACPRVTSCADLLIAAQQSVTLAGGFSWRV 120

QY 121 PLGRDLSQAFLDLANANLPAPFTTLPOLKDSFRNVLNRSSDLVALSGGHTFGKNQCRF 180
Db 121 PLGRDLSQAFLDLANANLPAPFTTLPOLKDSFRNVLNRSSDLVALSGGHTFGKNQCRF 180

QY 181 IMRLYNFNTGLPDPNTLNTYLTQLRGLCPNLGNLSALVDFDLRTPTTFDNKYVYNLEE 240
Db 181 IMRLYNFNTGLPDPNTLNTYLTQLRGLCPNLGNLSALVDFDLRTPTTFDNKYVYNLEE 240

QY 241 OKGLIQSQDELSPSPATDTIPLVRFSANSTQTFEAFVEMDRMGNTPLTGTGQIRL 300
Db 241 OKGLIQSQDELSPSPATDTIPLVRFSANSTQTFEAFVEMDRMGNTPLTGTGQIRL 300

QY 301 NCRVNSNS 309
Db 301 NCRVNSNS 309

RESULT 10
US-10-274-793-24
; Sequence 24, Application US/10274793
; Publication No. US20030153042A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY;
; APPLICANT: Frances S. ARNOLD
; APPLICANT: Zhanglin LIN
; TITLE OF INVENTION: Expression of Functional Eukaryotic
; TITLE OF INVENTION: proteins
; FILE REFERENCE: 3369/1E804-US3
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/654,493
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/538,591
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/247,232
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-274-793-24

Query Match 98.1%; Score 1577; DB 14; Length 309;
Best Local Similarity 98.4%; Pred. No. 1.5e-159;
Matches 304; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MQLPTFYDNCSPNVNIVRDTIVNELRSPRIAASILRLHFHDFCVNGCDASILLDNTT 60
Db 1 MQLPTFYDNCSPNVNIVRDTIVNELRSPRIAASILRLHFHDFCVNGCDASILLDNTT 60

QY 61 SFRTEKDAFGNANSARGFPVIDRMKAAVESACPRVTSCADLLIAAQQSVTLAGGFSWRV 120
Db 61 SFRTEKDAFGNANSARGFPVIDRMKAAVESACPRVTSCADLLIAAQQSVTLAGGFSWRV 120

QY 121 PLGRDLSQAFLDLANANLPAPFTTLPOLKDSFRNVLNRSSDLVALSGGHTFGKNQCRF 180
Db 121 PLGRDLSQAFLDLANANLPAPFTTLPOLKDSFRNVLNRSSDLVALSGGHTFGKNQCRF 180

QY 181 IMRLYNFNTGLPDPNTLNTYLTQLRGLCPNLGNLSALVDFDLRTPTTFDNKYVYNLEE 240
Db 181 IMRLYNFNTGLPDPNTLNTYLTQLRGLCPNLGNLSALVDFDLRTPTTFDNKYVYNLEE 240

QY 241 OKGLIQSQDELSPSPATDTIPLVRFSANSTQTFEAFVEMDRMGNTPLTGTGQIRL 300
Db 241 OKGLIQSQDELSPSPATDTIPLVRFSANSTQTFEAFVEMDRMGNTPLTGTGQIRL 300

QY 301 NCRVNSNS 309
Db 301 NCRVNSNS 309
```

```
; TITLE OF INVENTION: Expression of Functional Eukaryotic
; TITLE OF INVENTION: proteins
; FILE REFERENCE: 3369/1E804-US3
; CURRENT APPLICATION NUMBER: US/10/274,793
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/654,493
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/538,591
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/247,232
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 60/106,840
; PRIOR FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-274-793-24

Query Match 98.1%; Score 1577; DB 14; Length 309;
Best Local Similarity 98.4%; Pred. No. 1.5e-159;
Matches 304; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MQLPTFYDNCSPNVNIVRDTIVNELRSPRIAASILRLHFHDFCVNGCDASILLDNTT 60
Db 1 MQLPTFYDNCSPNVNIVRDTIVNELRSPRIAASILRLHFHDFCVNGCDASILLDNTT 60

QY 61 SFRTEKDAFGNANSARGFPVIDRMKAAVESACPRVTSCADLLIAAQQSVTLAGGFSWRV 120
Db 61 SFRTEKDAFGNANSARGFPVIDRMKAAVESACPRVTSCADLLIAAQQSVTLAGGFSWRV 120

QY 121 PLGRDLSQAFLDLANANLPAPFTTLPOLKDSFRNVLNRSSDLVALSGGHTFGKNQCRF 180
Db 121 PLGRDLSQAFLDLANANLPAPFTTLPOLKDSFRNVLNRSSDLVALSGGHTFGKNQCRF 180

QY 181 IMRLYNFNTGLPDPNTLNTYLTQLRGLCPNLGNLSALVDFDLRTPTTFDNKYVYNLEE 240
Db 181 IMRLYNFNTGLPDPNTLNTYLTQLRGLCPNLGNLSALVDFDLRTPTTFDNKYVYNLEE 240

QY 241 OKGLIQSQDELSPSPATDTIPLVRFSANSTQTFEAFVEMDRMGNTPLTGTGQIRL 300
Db 241 OKGLIQSQDELSPSPATDTIPLVRFSANSTQTFEAFVEMDRMGNTPLTGTGQIRL 300

QY 301 NCRVNSNS 309
Db 301 NCRVNSNS 309

RESULT 11
US-10-274-793-22
; Sequence 22, Application US/10274793
; Publication No. US20030153042A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY;
; APPLICANT: Frances S. ARNOLD
; APPLICANT: Zhanglin LIN
; TITLE OF INVENTION: Expression of Functional Eukaryotic
; TITLE OF INVENTION: proteins
; FILE REFERENCE: 3369/1E804-US3
; CURRENT APPLICATION NUMBER: US/10/274,793
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/654,493
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 09/538,591
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 09/247,232
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: US 60/094,403
; PRIOR FILING DATE: 1998-07-28
```

; PRIOR APPLICATION NUMBER: US 60/106,840
 ; PRIOR FILING DATE: 1998-11-03
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 22
 ; LENGTH: 309
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 US-10-274-793-22

Query Match 97.8%; Score 1572; DB 14; Length 309;
 Best Local Similarity 93.1%; Pred. No. 5.3e-89;
 Matches 303; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY	1	MQLPTFYDNCSPVNSIVRDTIIVNELRSDPRTIAASILRLHFDHCFVNGCDASILLDNTT	60
DB	1	MQLPTFYDNCSPVNSIVRDTIIVNELRSDPRTIAASILRLHFDHCFVNGCDASILLDNTT	60
QY	61	SPTEKDAFGNANSARGFPVIDRMKAIVESACPTVSCADLLTIAAQSVTLAGGFSWRV	120
DB	61	SPTEKDAFGNANSARGFPVIDRMKAIVESACPTVSCADLLTIAAQSVTLAGGFSWRV	120
QY	121	PGRDLSLOAFDLANANLPAPFTLPOLKDSFRNVGLNRSSDLVALSGHTEGKNOCRF	180
DB	121	PGRDLSLOAFDLANANLPAPFTLPOLKDSFRNVGLNRSSDLVALSGHTEGKNOCRF	180
QY	181	IMRLYNFNTGLPDTLNTTYLQTLRGLCPNGNLSALVDFDLRTPTIFONKYYVNL	240
DB	181	IMRLYNFNTGLPDTLNTTYLQTLRGLCPNGNLSALVDFDLRTPTIFONKYYVNL	240
QY	241	QKGLIQSDQELFSSPDATDITPLVRFSANSTQTFNFAVEAMDRMGNTITLTGQOIRL	300
DB	241	QKGLIQSDQELFSSPDATDITPLVRFSANSTQTFNFAVEAMDRMGNTITLTGQOIRL	300
QY	301	NCRVNSNS 309	
DB	301	NCREVNSNS 309	

RESULT 12
 US-10-424-595-278555
 ; Sequence 278555, Application US/10424599
 ; Publication No. US20040031072A1

; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 278555
 ; LENGTH: 352
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_93558C.1.pap
 US-10-424-599-278555

Query Match 57.3%; Score 921; DB 12; Length 352;
 Best Local Similarity 59.1%; Pred. No. 2.2e-89;
 Matches 182; Conservative 40; Mismatches 84; Indels 2; Gaps 2;

QY	2	QLTPTFYDNCSPVNSIVRDTIIVNELRSDPRTIAASILRLHFDHCFVNGCDASILLDNTTS	61
DB	27	QLTPTFYRETCPNLPFIVFGIFDASFTDPRIGASILRLHFDHCFVNGCDASILLDNTT	86
QY	62	FTEKDAFGNANSARGFPVIDRMKAIVESACPTVSCADLLTIAAQSVTLAGGFSWRV	121
DB	87	ISEQDALPNIINSIRGLDVVNDIKTAVKACPGWSCADILITLASEITSSVLGGGPDWKVP	146

QY	122	LGRDLSLOAFDLANANLPAPFTLPOLKDSFRNVGLNRSSDLVALSGHTEGKNOCRF	181
DB	147	LGRDLSLTANRTLANQNLPAPFENLTQKASFAVOGLN-TLTLVTLSSGHTFGARCSFT	205
QY	182	MDRLYNFNTGLPDTLNTTYLQTLRGLCPNGNLSALVDFDLRTPTIFONKYYVNL	241
DB	206	INRLYNFNTGNPDPTLNTTYLQTLRGLCPNGNLSALVDFDLRTPTIFONKYYVNL	265
QY	242	KGLIQSDQELFSSPDATDITPLVRFSANSTQTFNFAVEAMDRMGNTITLTGQOIRL	301
DB	266	NGLLQSDQELFSTPGA-DTIPVNSFSSMONTFFSNFRVSMIKMGNIGVLTGDEGEIRLQ	324
QY	302	CRVNSNS 309	
DB	325	CNEVNGDS 332	

RESULT 13

US-10-424-599-250229
 ; Sequence 250229, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 250229
 ; LENGTH: 354
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_67986C.1.pap
 US-10-424-599-250229

Query Match 56.7%; Score 912.5; DB 12; Length 354;
 Best Local Similarity 57.5%; Pred. No. 1.8e-88;
 Matches 177; Conservative 50; Mismatches 78; Indels 3; Gaps 3;

QY	2	QLTPTFYDNCSPVNSIVRDTIIVNELRSDPRTIAASILRLHFDHCFVNGCDASILLDNTTS	61
DB	28	QLDPSFYRTCTCRVHSIVREVRNVSKXPRMLASILRLHFDHCFVNGCDASILLDNTAT	87
QY	62	FTEKDAFGNANSARGFPVIDRMKAIVESACPTVSCADLLTIAAQSVTLAGGFSWRV	121
DB	88	ISEQDALPNIINSIRGLDVVNDIKTAVKACPGWSCADILITLASEITSSVLGGGPDWKVP	147
QY	122	LGRDLSLOAFDLANANLPAPFTLPOLKDSFRNVGLNRSSDLVALSGHTEGKNOCRF	181
DB	148	LGRDLSLTANRTLANQNLPAPFENLTQKASFAVOGLD-TTDLVLSGAHTFGARCNFI	206
QY	182	MDRLYNFNTGLPDTLNTTYLQTLRGLCPNGNLSALVDFDLRTPTIFONKYYVNL	241
DB	207	LDRLYNFNTGTXPDPTLDTTYLQQLRQICP-NGGPNLNVFDPVTPDKIDRVVFSNLQVK	265
QY	242	KGLIQSDQELFSSPDATDITPLVRFSANSTQTFNFAVEAMDRMGNTITLTGQOIRL	301
DB	266	KGLIQSDQELFSTPGA-DTIPVNSFSSDQKVFVDFDAFEASIMKMGNIGVLTGKKEIRKH	324
QY	302	CRVNSNS 309	
DB	325	CNEVAKKS 332	

RESULT 14

US-10-424-599-251628
 ; Sequence 251628, Application US/10424599

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 18:49:16 ; Search time 11.1649 Seconds
(without alignments)
1428.803 Million cell updates/sec

Title: US-09-246-451A-17

Perfect score: 1608

Sequence: 1 MQLTPTFYDNCSPVNSIVR.....PLFTGTCQIRNCRVNSNS 309

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep:*

2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep:*

3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep:*

4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep:*

5: /cgn2_6/ptodata/2/iaa/PTUS.COMB.pep:*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1462	90.9	349	1	US-08-190-029A-10
2	1462	90.9	349	2	US-08-462-695-10
3	921	57.3	352	4	US-09-207-914-22
4	781.5	48.6	324	3	US-09-365-150-2
5	768.5	47.8	333	4	US-09-097-319A-2
6	727.5	45.2	318	4	US-09-615-192A-389
7	722.5	44.9	323	4	US-09-615-192A-395
8	719.5	44.7	315	4	US-09-615-192A-378
9	646.5	40.2	287	4	US-09-615-192A-387
10	594.5	37.0	324	2	US-08-671-320-11
11	594.5	37.0	324	2	US-08-368-577-11
12	594.5	37.0	324	4	US-09-207-914-11
13	591.5	36.8	324	2	US-08-671-320-13
14	591.5	36.8	324	2	US-08-668-577-13
15	591.5	36.8	324	4	US-09-207-914-13
16	590	36.7	266	4	US-09-615-192A-394
17	587.5	36.5	351	4	US-09-615-192A-397
18	566	35.2	308	4	US-09-615-192A-381
19	566	35.2	313	2	US-08-671-320-15
20	566	35.2	313	2	US-08-668-577-15
21	566	35.2	313	4	US-09-207-914-15
22	565	35.1	313	2	US-08-671-320-17
23	565	35.1	313	2	US-08-668-577-17
24	565	35.1	313	4	US-09-207-914-17
25	503	31.3	203	4	US-09-615-192A-396
26	458.5	28.5	201	4	US-09-615-192A-39-
27	433.5	27.0	208	4	US-09-615-192A-385

28 372.5 23.2 138 4 US-09-615-192A-384 Sequence 384, App

29 365.5 22.7 161 4 US-09-615-192A-388 Sequence 388, App

30 343 21.3 143 4 US-09-615-192A-401 Sequence 401, App

31 331 20.6 157 4 US-09-615-192A-399 Sequence 399, App

32 331 20.6 170 4 US-09-615-192A-332 Sequence 332, App

33 312.5 19.4 179 4 US-09-615-192A-376 Sequence 376, App

34 298.5 18.6 121 1 US-08-190-029A-12 Sequence 12, Appl

35 298.5 18.6 121 2 US-08-462-695-12 Sequence 12, Appl

36 288 17.9 202 4 US-09-615-192A-386 Sequence 386, App

37 278.5 17.3 120 4 US-09-615-192A-393 Sequence 393, App

38 268.5 16.7 117 4 US-09-615-192A-400 Sequence 400, App

39 244.5 15.2 114 4 US-09-615-192A-347 Sequence 347, App

40 241 15.0 111 4 US-09-615-192A-379 Sequence 379, App

41 237.5 14.8 95 4 US-09-615-192A-390 Sequence 390, App

42 235.5 14.6 118 4 US-09-615-192A-333 Sequence 333, App

43 200 12.4 120 4 US-09-615-192A-392 Sequence 392, App

44 187 11.6 103 4 US-09-615-192A-398 Sequence 398, App

45 121.5 7.6 365 4 US-09-748-264A-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-08-190-029A-10

; Sequence 10, Application US/08190029A

; Patent No. 5736363

; GENERAL INFORMATION:

; APPLICANT: EDWARDS, Richard Mark

; APPLICANT: BANDEN, Lindsey

; TITLE OF INVENTION: IGF-II ANALOGUES

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ALLEGRETTI & WITCOFF, LTD.

; STREET: 10 S. WACKER DRIVE, SUITE 3000

; CITY: CHICAGO

; STATE: ILLINOIS

; COUNTRY: U.S.A.

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/190,029A

; FILING DATE: 28-FEB-1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/GB92/01389

; FILING DATE: 27-JUL-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9202401.7

; FILING DATE: 05-FEB-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9116325.3

; FILING DATE: 29-JUL-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: JOHN J. McDONNELL

; REGISTRATION NUMBER: 26,949

; REFERENCE/DOCKET NUMBER: 94,062

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312-715-1000

; TELEFAX: 312-715-1234

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 349 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-190-029A-10

Query March 90.9%; Score 1462; DB 1; Length 349;

Wed Apr 7 09:15:32 2004

MOLECULE TYPE: protein
US-08-462-695-10

Query Match 90.9%; Score 1462; DB 2; Length 349;
Best Local Similarity 99.6%; Pred. No. 3.8e-151;
Matches 281; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQLTPTFYDNCSPNVNIVRDTIVNELRSDPRTIAASILRLHFDHCFVNGCDASILLDNTT 60
DB 1 MQLTPTFYDNCSPNVNIVRDTIVNELRSDPRTIAASILRLHFDHCFVNGCDASILLDNTT 60

QY 61 SPTKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGSPWRV 120
DB 61 SPTKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGSPWRV 120

QY 121 PLGRDLSLOAFDLANANLPAPFTTLPOLKDSFRNVGLNRSSDLVALSGHTEGKNQCRF 180
DB 121 PLGRDLSLOAFDLANANLPAPFTTLPOLKDSFRNVGLNRSSDLVALSGHTEGKNQCRF 180

QY 181 IMRLYNFSNTGLPDPPTINTTYLQTLRGCLPLNGLSALVDFDLRTPTTFDNKYVNL 240
DB 181 IMRLYNFSNTGLPDPPTINTTYLQTLRGCLPLNGLSALVDFDLRTPTTFDNKYVNL 240

QY 241 QKGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNFAFVEAM 282
DB 241 QKGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNFAFVEAM 282

RESULT 3
US-09-207-914-22
; Sequence 22, Application US/09207914A
; Patent No. 6586583
; GENERAL INFORMATION:
; APPLICANT: Vierling Jr., Richard A.
; TITLE OF INVENTION: A Soybean Peroxidase Gene Family and an Assay for
; TITLE OF INVENTION: Detecting Soybean Peroxidase Activity
; FILE REFERENCE: Soybean Peroxidase Gene Family
; CURRENT APPLICATION NUMBER: US/09/207,914A
; CURRENT FILING DATE: 1998-12-09
; EARLIER APPLICATION NUMBER: US 08/868,577
; EARLIER FILING DATE: 1997-06-04
; EARLIER APPLICATION NUMBER: US 08/671,320
; EARLIER FILING DATE: 1995-10-27
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 22
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Glycine max
US-09-207-914-22

Query Match 57.3%; Score 921; DB 4; Length 352;
Best Local Similarity 59.1%; Pred. No. 5e-92;
Matches 182; Conservative 40; Mismatches 84; Indels 2; Gaps 2;

QY 2 QLTFTFYDNCSPNVNIVRDTIVNELRSDPRTIAASILRLHFDHCFVNGCDASILLDNTT 61
DB 27 QLTFTFYRETCPNLFPIVGFIDASFDDPRIGLSMLRHLHFDHCFVQCGDGSVILNNTDT 86

QY 62 PRTKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGSPWRV 121
DB 87 TESEODALPNINSIRGLDVNDIKTAVENSCPTVSCADLLTIAAAEIASVLGGGPGWVP 146

QY 122 LGRRDLSLOAFDLANANLPAPFTTLPOLKDSFRNVGLNRSSDLVALSGHTEGKNQCRF 181
DB 147 LGRRDLSLTANRTLANONLPAPPENLTQLKASFAVQGLN--TLDLVTLGSGHTEGACSTF 205

QY 162 MDRLYNFSNTGLPDPPTINTTYLQTLRGCLPLNGLSALVDFDLRTPTTFDNKYVNL 241
DB 206 INRLYNFSNTGNPDPTINTTYLQTLRGCLPLNGLSALVDFDLRTPTTFDNKYVNL 255

QY 242 KGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNFAFVEAMDMGNITLTGTGQIRLN 301

Best Local Similarity 99.6%; Pred. No. 3.8e-151;
Matches 281; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQLTPTFYDNCSPNVNIVRDTIVNELRSDPRTIAASILRLHFDHCFVNGCDASILLDNTT 60
DB 1 MQLTPTFYDNCSPNVNIVRDTIVNELRSDPRTIAASILRLHFDHCFVNGCDASILLDNTT 60

QY 61 SPTKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGSPWRV 120
DB 61 SPTKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGSPWRV 120

QY 121 PLGRDLSLOAFDLANANLPAPFTTLPOLKDSFRNVGLNRSSDLVALSGHTEGKNQCRF 180
DB 121 PLGRDLSLOAFDLANANLPAPFTTLPOLKDSFRNVGLNRSSDLVALSGHTEGKNQCRF 180

QY 181 IMRLYNFSNTGLPDPPTINTTYLQTLRGCLPLNGLSALVDFDLRTPTTFDNKYVNL 240
DB 181 IMRLYNFSNTGLPDPPTINTTYLQTLRGCLPLNGLSALVDFDLRTPTTFDNKYVNL 240

QY 241 QKGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNFAFVEAM 282
DB 241 QKGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNFAFVEAM 282

RESULT 2
US-08-462-695-10
; Sequence 10, Application US/08462695
; Patent No. 5854025
; GENERAL INFORMATION:
; APPLICANT: EDWARDS, Richard Mark
; APPLICANT: BAWDEN, Lindsey
; TITLE OF INVENTION: IGF-II ANALOGUES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESS: BANNER & ALLEGRETTI, LTD.
; STREET: 10 S. WACKER DRIVE, SUITE 3000
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: U.S.A.
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,695
; FILING DATE: 5-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/190,029
; FILING DATE: 28-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/01389
; FILING DATE: 27-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9202401.7
; FILING DATE: 05-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9116325.3
; FILING DATE: 29-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: JOHN J. McDONNELL
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 94,062-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 349 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

Db 266 NGLQSDQELFSTPGA-DTPIVNSFSSNQNTFFSNFRVSMIRKMGNIGVLTGDEGEIRLQ 324
QY 302 CRVANSNS 309
Db 325 CNFVNGDS 332

RESULT 4

US-09-365-150-2
; Sequence 2, Application US/09365150
; Patent No. 6278041
; GENERAL INFORMATION:
; APPLICANT: Lagrimini, Mark
; APPLICANT: Desai, Nalini
; TITLE OF INVENTION: No. 6278041el Peroxidase Gene Sequences
; CURRENT APPLICATION NUMBER: JS/09/365,150
; FILE REFERENCE: S-3-081P1
; CURRENT FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Kocotiana tomentosiformis
US-09-365-150-2

Query Match 48.6%; Score 781.5; DB 3; Length 324;
Best Local Similarity 50.3%; Pred. No. 7.7e-77;
Matches 154; Conservative 59; Mismatches 88; Indels 5; Gaps 5;
QY 2 QUTPTFYNSCNVSNVVRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLNTTSF 61
Db 23 QLSATFYDSTCPNVTSEVRGVMQDQRTDARAGAKIIRLHFHDCFVNGCDGSIILLD-IDG 81
QY 62 FTEKDAFNGANSARGFVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGSPWRVP 121
Db 82 TOTEDKADPNV-GAGGFVIDDITKALENVCPGWSCADILSLASIGVALAEGPSQWL 140
QY 122 LGRDLSQAFDLANAKLPAPFTTLPQKDSFRNVGLNRSSDLVALSGGHTFGKNCRTI 181
Db 141 FGKNSLTANRSEANS-ESPETPAVMTPLTNKGM-DTLVAGSGAHTFGARCGTF 199
QY 182 MDRLNFSNTGLPDTNTTY-QTLRGLCPNGNL-SALVDFDLRPTTTFIDNKYYVNLLE 240
Db 200 EQRLENFSGNPDPTVATFLQTLQICPOGNGNGTFTNLIDSTPNDFDNDYFNLQN 259
QY 241 QKGLTQSDQLFSSPDATDTIPLVRSFANSTOTFENAFVEMDRGNITPLTGTQCIIRL 300
Db 260 NQSLQTDQELFST-SGSATIAVVRVYAGSQTFDDFVSSMIKLGNSPLGTNGEIRT 318
QY 301 MCRVYN 306
Db 319 DCXRVN 324

RESULT 5

US-09-097-319A-2
; Sequence 2, Application US/09097319A
; Patent No. 6384207
; GENERAL INFORMATION:
; APPLICANT: Ainley, Michael
; APPLICANT: Armstrong, Katherine
; APPLICANT: Belmar, Scott
; APPLICANT: Folkerts, Otto
; APPLICANT: Hopkins, Nicole
; APPLICANT: Menke, Michael A.
; APPLICANT: Paredy, Dayakar
; APPLICANT: Petolino, Joseph F.
; APPLICANT: Smith, Kelley
; APPLICANT: Woosley, Aaron
; TITLE OF INVENTION: Regulatory Sequences for Transgenic Plants
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DowElanco Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/097,319A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317 337 4816
; TELEFAX: 317 337 4847
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 333 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-097-319A-2
Query Match 47.8%; Score 768.5; DB 4; Length 333;
Best Local Similarity 49.7%; Pred. No. 2.1e-75;
Matches 152; Conservative 52; Mismatches 99; Indels 3; Gaps 2;
QY 3 LPTFTVNSCNVSNVVRDTIVNELRSDPRIAASILRLHFHDCFVNGCDASILLNTTSF 62
Db 30 LFPQFYDHSCKPAXE-VQSIQAQVAKETRMASLVELHFDHCFWKGCDASVLLDNSSI 89
QY 63 RTEKDAFNGANSARGFVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGSPWRVPL 122
Db 90 VSEKGSNPNSILRGEVIDQIKALEACPTVSCADIVALLAARUSTALVGGPYWDVPL 149
QY 123 GRDLSQAFDLANAKLPAPFTTLPQKDSFRNVGLNRSSDLVALSGGHTFGKNCQRFIM 182
Db 150 GRDLSGASIQGSNDIPAPNNTLPTIITKFKQGLN-VVDVVALSGGHTIGMSRCTSF 208
QY 183 ERLNFSNTGLPDTNTTYLQTLRGLCPNGNL-SALVDFDLRPTTTFIDNKYYVNLLEQK 242
Db 209 QRLNQTGNGMADSTUDVSTAANKLRQCPRSGGNNLFLDLFITPAKFDNFYKLNLAGX 268
QY 243 GLTQSDQLFSSPDATDTIPLVRSFANSTOTFENAFVEMDRGNITPLTGTQCIIRLNC 302
Db 269 GLLSDEILLTK--SAETAALVKAYADYNLFFQHPAQSMVNMGNISPLTSGQGEIRKNC 326
QY 303 RVVNSN 308
Db 327 RLINND 332

RESULT 6

US-09-615-192A-389
; Sequence 389, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003C4U
; CURRENT APPLICATION NUMBER: US/09/615,192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11

```

; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 389
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-615-192A-389

Query Match      45.2%; Score 727.5; DB 4; Length 318;
Best Local Similarity 49.5%; Pred. No. 5.9e-71;
Matches 151; Conservative 48; Mismatches 95; Indels 11; Gaps 4;

QY 2 QLTPTFYDNCNVNIVRDTIVNELRSDPRTVSCADLLTIAAQSVTLAGGFSWRVP 61
DB 25 QLSSTFYAKSCPRLSIVKAVKQVAVKXKRGASIVRLHFHDCFVNGCDGSLILDDNAT 84
QY 62 FRTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGFSWRVP 121
DB 85 FTGEXTAGFNANSARGFEVIDSIKTQVEAAACGWSVCADILITIAARDSIVELQGPWTVM 144
QY 122 LGRDLSQAFDLANANLPAPFTLPQKDSFRNVGLNRSSDLVALSGGHTFGKNCQCFI 181
DB 145 LGRDSTTASAAANNIPSPASSLTLSSFOAEGLS-TKDLVALSGAHTTQGRCAFF 203
QY 182 MDRLYNFSNTGLPDTLNTTYLTQTLRGLCPLNGLSALVDFDLRTPTTFIDNKYYVLEBQ 241
DB 204 RTRIYNEN-----INRAFTSVKANCPSAGGSNLSPLDAVTSITFDNKYYSNLKIQ 256
QY 242 KGLISDOELFSSPDATDIPIVRFSFANSTQTFNFAFVEAMDRMGNITPLTGTQOIRN 301
DB 257 KGLHSDQOLFNG-GSTDS--QVTYSSNQNSFFIDFTAAVYKMGNSPLTGTGQIRKN 313
QY 302 CRVUN 306
DB 314 CRKSN 318

RESULT 7
US-09-615-192A-395
; Sequence 395, Application US/096:5192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1003C4U
; CURRENT APPLICATION NUMBER: US 09/615,192A
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-03
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 395
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-615-192A-395

Query Match      44.9%; Score 722.5; DB 4; Length 323;
Best Local Similarity 49.0%; Pred. No. 2.1e-70;
Matches 152; Conservative 49; Mismatches 90; Indels 19; Gaps 6;

QY 2 QLTPTFYDNCNVNIVRDTIVNELRSDPRTVSCADLLTIAAQSVTLAGGFSWRVP 61
DB 28 QLSSTFYKSCPTALSIVKAVKQVAVKXKRGASLRLHFHDCFVNGCDGSLILDDST 87

; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 389
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-615-192A-389

Query Match      45.2%; Score 727.5; DB 4; Length 318;
Best Local Similarity 49.5%; Pred. No. 5.9e-71;
Matches 151; Conservative 48; Mismatches 95; Indels 11; Gaps 4;

QY 2 QLTPTFYDNCNVNIVRDTIVNELRSDPRTVSCADLLTIAAQSVTLAGGFSWRVP 61
DB 25 QLSSTFYAKSCPRLSIVKAVKQVAVKXKRGASIVRLHFHDCFVNGCDGSLILDDNAT 84
QY 62 FRTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGFSWRVP 121
DB 85 FTGEXTAGFNANSARGFEVIDSIKTQVEAAACGWSVCADILITIAARDSIVELQGPWTVM 144
QY 122 LGRDLSQAFDLANANLPAPFTLPQKDSFRNVGLNRSSDLVALSGGHTFGKNCQCFI 181
DB 145 LGRDSTTASAAANNIPSPASSLTLSSFOAEGLS-TKDLVALSGAHTTQGRCAFF 203
QY 182 MDRLYNFSNTGLPDTLNTTYLTQTLRGLCPLNGLSALVDFDLRTPTTFIDNKYYVLEBQ 241
DB 204 RTRIYNEN-----INRAFTSVKANCPSAGGSNLSPLDAVTSITFDNKYYSNLKIQ 256
QY 242 KGLISDOELFSSPDATDIPIVRFSFANSTQTFNFAFVEAMDRMGNITPLTGTQOIRN 301
DB 257 KGLHSDQOLFNG-GSTDS--QVTYSSNQNSFFIDFTAAVYKMGNSPLTGTGQIRKN 313
QY 302 CRVUN 306
DB 314 CRKSN 318

RESULT 8
US-09-615-192A-378
; Sequence 378, Application US/096:5192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1003C4U
; CURRENT APPLICATION NUMBER: US 09/615,192A
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-03
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 378
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-09-615-192A-378

Query Match      44.7%; Score 719.5; DB 4; Length 315;
Best Local Similarity 49.2%; Pred. No. 4.4e-70;
Matches 151; Conservative 46; Mismatches 95; Indels 15; Gaps 4;

QY 2 QLTPTFYDNCNVNIVRDTIVNELRSDPRTVSCADLLTIAAQSVTLAGGFSWRVP 61
DB 22 KLSPSHYOSTCPKALSIVRAGVAKAIKNETRTGSLRLHFHDCFVNGCDASILLDDTFS 81
QY 62 FRTEKDAFGNANSARGFPVIDRMKAAVESACPTVSCADLLTIAAQSVTLAGGFSWRVP 121
DB 82 FVGEKTAAPNNNSVGEFVDRKASLEKECEGVVSCADIVALAARDSVVLHGGFSWTYS 141
QY 122 LGRDLSQAFDLANANLPAPFTLPQKDSFRNVGLNRSSDLVALSGGHTFGKNCQCFI 181
DB 142 LGRKDSITASRSLANTSIPPTSNLSALITSAQAQGLS-VKMMVALSGSHTTCLARCTSF 200
QY 182 MDRLYNFSNTGLPDTLNTTYLTQTLRGLCPLNGLSALVDFDLRTPTTFIDNKYYVLEBQ 241
DB 201 RRRIFYDNCN-----IDTSFAHKLQKICPRIGNDSVLQRLDIQTPTTFIDNLYHNLQK 253
QY 242 KGLISDOELF--GSPDATDIPIVRFSFANSTQTFNFAFVEAMDRMGNITPLTGTQOIR 299
DB 254 KGLHSDQOLFNGSSVDS-----LVKKYACTGKFFRFKAKWIKMSEIKPKGSGNQIR 308
QY 300 LNCRVUN 306

```

Db 309 KNCRKVN 315

RESULT 9

US-09-615-192A-387
 ; Sequence 387, Application US/09615192A
 ; Patent No. 6410718
 ; GENERAL INFORMATION:
 ; APPLICANT: Bloksberg, Leonard N.
 ; APPLICANT: Havukala, Ilkka
 ; TITLE OF INVENTION: Materials and Methods for the
 ; TITLE OF INVENTION: Modification of Plant Lignin Content
 ; FILE REFERENCE: 11000.1003c4U
 ; CURRENT APPLICATION NUMBER: US/09/615.192A
 ; CURRENT FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 08/975,316
 ; PRIOR FILING DATE: 1997-11-21
 ; PRIOR APPLICATION NUMBER: US 08/713,000
 ; PRIOR FILING DATE: 1996-09-11
 ; PRIOR APPLICATION NUMBER: US 09/163,789
 ; PRIOR FILING DATE: 1998-10-09
 ; NUMBER OF SEQ ID NOS: 405
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 387
 ; LENGTH: 287
 ; TYPE: PRT
 ; ORGANISM: Pinus radiata
 US-09-615-192A-387

Query Match 40.2%; Score 646.5; DB 4; Length 287;
 Best Local Similarity 48.6%; Pred. No. 3.6e-62;
 Matches 135; Conservative 46; Mismatches 78; Indels 19; Gaps 6;
 QY 2 QLTPTFYDNCSPNVNIVRDTIVNELRSDPRIAASILRLHFDHCFVNGCDASILLNDNTTS 61
 Db 24 KLSITFDKCPKLSIVQGVKQAVAKKMGASLLRLHFDHCFVNGCDGSVLLDNTT 83
 QY 62 FTEKDAFGNANSARGFPVIDRMKAIVESACPRVTS CADLTLIAAQSVTLAGGPSWRVP 121
 Db 84 FTEKDALPNNNSARGFEVIDSIKQLENACTGVVSCADILTIARDSVWVLGGPSWKVM 143
 QY 122 LGRDLSQAFDLANANLPAPFFTLPOLKDSFRNVGLNRSSDLVALSGGHTFGKQCRFI 181
 Db 144 LGRDSTASISGANNIPPTSLNLTLSLQFQAGLS-TKEMVALSGGHTFGQAQCKNF 202
 QY 182 MDRLYNPSNTGLPDTLNTVLTQTLRGL-NGNLSALVDFDLRPTTFDNKYV 236
 Db 203 RAIYN-----DTNIDTITATSLRCKSPSTTGSGDSNLSPL---DYTTFTVFDKNYY 252
 QY 237 NLEEQGLIQSDQELFSSPDATDTIPLVRGFANSTOTF 274
 Db 253 NLKSKRGLLHSDQELFNG-GSTDS--HVTKYASNQNTF 287

RESULT 10

US-08-671-320-11
 ; Sequence 11, Application US/08671320
 ; Patent No. 5840558
 ; GENERAL INFORMATION:
 ; APPLICANT: VIERLING JR, RICHARD A
 ; TITLE OF INVENTION: A SOYBEAN PEROXIDASE GENE FAMILY AND AN
 ; TITLE OF INVENTION: ASSAY FOR DETECTING SOYBEAN PEROXIDASE ACTIVITY
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: 1125 SO. 103RD STREET
 ; STREET: SUITE 330
 ; CITY: OMAHA
 ; STATE: NE
 ; COUNTRY: US
 ; ZIP: 68124-1076
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Versior #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/671,320
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: JONDLE, ROBERT J.
 REGISTRATION NUMBER: 33,915
 REFERENCE/DOCKET NUMBER: 1227-001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 402-398-9000
 TELEFAX: 402-398-9005
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 324 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-671-320-11

Query Match 37.0%; Score 594.5; DB 2; Length 324;
 Best Local Similarity 39.9%; Pred. No. 2.1e-56;
 Matches 123; Conservative 63; Mismatches 115; Indels 7; Gaps 5;
 QY 2 QLTPTFYDNCSPNVNIVRDTIVNELRSDPRIAASILRLHFDHCFVNGCDASILLNDNTTS 61
 Db 22 QLQGFYANSCPKAQIIVLFVHDHINAPSLAAALIRMHFDHCFVNGCDASVLLNSTN 81
 QY 62 FTEKDAFGNANSARGFPVIDRMKAIVESACPRVTS CADLTLIAAQSVTLAGGPSWRVP 121
 Db 82 -QAEKNAEPNL-TVRGFDIFDIKSLVEAECPGVVS CADILTLAARDTIVATGPGFWKVP 139
 QY 122 LGRDLSQAFDLANANLPAPFFTLPOLKDSFRNVGLNRSSDLVALSGGHTFGKQCRFI 181
 Db 140 TGRDGVVSNLTARNNIPAPSSNFTTLQTLFANQGLD-LKDLVLSGHTFGIAHCSSL 198
 QY 182 MDRLYNPSNTGLPDTLNTVLTQTLRGL-CPNLGNLSALVDFDLRPTTFDNKYV 239
 Db 199 SNRLNFNTGKGDQDPSLDSEYAAALKAFKCDLNLKNTTKIEMDPGSRKTFDLSYSHVI 258
 QY 240 EOKGLIQSDQELFSSPDATDTIPLVRGFANSTOTFENAFVEMBMGNITPLTGTGQOIR 299
 Db 259 KRGLEFSDAALLT-NSVTKAQIIQLLEGSVENFFAEFATSIEXMGRINVKGTGEIR 316
 QY 300 LNCRVVNS 307
 Db 317 KHCAPINS 324

RESULT 11
 US-08-868-577-11
 ; Sequence 11, Application US/08868577
 ; Patent No. 5866695
 ; GENERAL INFORMATION:
 ; APPLICANT: Vierling Jr., Richard A
 ; TITLE OF INVENTION: A SOYBEAN PEROXIDASE GENE FAMILY AND AN
 ; TITLE OF INVENTION: ASSAY FOR DETECTING SOYBEAN PEROXIDASE ACTIVITY
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Rotwell, Figg, Ernst & Kurz
 ; STREET: 555 13th Street NW, Suite 701 East
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20004
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WordPerfect 6.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/868,577

```

/ FILING DATE: 04-JUN-1997
/ CLASSIFICATION: 536
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Jondle, Robert J.
/ REGISTRATION NUMBER: 33,915
/ REFERENCE/DOCKET NUMBER: N1227-003
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 402-333-1550
/ TELEFAX: 402-333-1510
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 324 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-868-577-11

Query Match 37.0%; Score 594.5; DB 2; Length 324;
Best Local Similarity 39.9%; Pred. No. 2.1e-56;
Matches 123; Conservative 63; Mismatches 115; Indels 7; Gaps 5;

QY 2 QLTPTFYDNCSPVNSIVRDTIVNELRSDPRIAASILRLHFHDFCVNGCDASILLNTTS 61
Db 22 QLQGFYANSCPAEQIVLKFVHDHINAPSLAALIRMHFDFCVRGCDASVLLNSTN 81

QY 62 ERTEKDAFGNANSARGPPVIDRMKAAVESACPRVTWSCADLLITIAAQOSVTLGAGPSWRVP 121
Db 82 -QAEKNAPPNL-TVRGDFDIDRIKSLVEAECPGVWSCADILTLAARDTIIVATGGFWKVP 139

QY 122 LGRDLSQAFDLIANANLPAFFTLQPKDSFRVNGLSRSDLVALSGGHTFGKNQCRFI 181
Db 140 TGRDGSVNSLTERARNIPAPSNFTTLQTFANQGLD-LKDIVLLSGAHTIGIAHCSSL 198

QY 182 MDRLYNESNTGLPDPNTLTYLQTLRGL--CPLNGNLSALVDFDLRPTTIIDNKYYNLE 239
Db 199 SNRLNFTGKGQDQSLDSEYANLAKFKCTDLNKNTTKIEMDPGSRKTFDLSYSHVI 258

QY 240 EOKGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNFAVEMDRMGNTIPLTGTQGOIR 299
Db 259 KRGLFESDAALLT--NSVTKAQIIQLLEGSVENFFAEFATSIKMGRIINVKTGTGEIR 316

QY 300 LNCRVVNS 307
Db 317 KHCAFNS 324

RESULT 12
US-09-207-914-11
/ Sequence 11, Application US/09207914A
/ Patent No. 6586583
/ GENERAL INFORMATION:
/ APPLICANT: Vierling Jr., Richard A.
/ TITLE OF INVENTION: A Soybean Peroxidase Gene Family and an Assay for
/ FILE REFERENCE: Detecting Soybean Peroxidase Activity
/ CURRENT APPLICATION NUMBER: US/09/207,914A
/ CURRENT FILING DATE: 1998-12-09
/ EARLIER APPLICATION NUMBER: US 08/868,577
/ EARLIER FILING DATE: 1997-06-04
/ EARLIER APPLICATION NUMBER: US 08/671,320
/ EARLIER FILING DATE: 1995-10-27
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: Patent In Ver. 2.0
/ SEQ ID NO 11
/ LENGTH: 324
/ TYPE: PRT
/ ORGANISM: Glycine max
/ US-09-207-914-11

Query Match 37.0%; Score 594.5; DB 4; Length 324;
Best Local Similarity 39.9%; Pred. No. 2.1e-56;
Matches 123; Conservative 63; Mismatches 115; Indels 7; Gaps 5;

QY 2 QLTPTFYDNCSPVNSIVRDTIVNELRSDPRIAASILRLHFHDFCVNGCDASILLNTTS 61
Db 22 QLQGFYANSCPAEQIVLKFVHDHINAPSLAALIRMHFDFCVRGCDASVLLNSTN 81

QY 62 ERTEKDAFGNANSARGPPVIDRMKAAVESACPRVTWSCADLLITIAAQOSVTLGAGPSWRVP 121
Db 82 -QAEKNAPPNL-TVRGDFDIDRIKSLVEAECPGVWSCADILTLAARDTIIVATGGFWKVP 139

QY 122 LGRDLSQAFDLIANANLPAFFTLQPKDSFRVNGLSRSDLVALSGGHTFGKNQCRFI 181
Db 140 TGRDGSVNSLTERARNIPAPSNFTTLQTFANQGLD-LKDIVLLSGAHTIGIAHCSSL 198

QY 182 MDRLYNESNTGLPDPNTLTYLQTLRGL--CPLNGNLSALVDFDLRPTTIIDNKYYNLE 239
Db 199 SNRLNFTGKGQDQSLDSEYANLAKFKCTDLNKNTTKIEMDPGSRKTFDLSYSHVI 258

QY 240 EOKGLIQSDQELFSSPDATDTIPLVRSFANSTQTFNFAVEMDRMGNTIPLTGTQGOIR 299
Db 259 KRGLFESDAALLT--NSVTKAQIIQLLEGSVENFFAEFATSIKMGRIINVKTGTGEIR 316

QY 300 LNCRVVNS 307
Db 317 KHCAFNS 324

RESULT 13
US-08-671-320-13
/ Sequence 13, Application US/08671320
/ Patent No. 5840558
/ GENERAL INFORMATION:
/ APPLICANT: Vierling Jr., Richard A.
/ TITLE OF INVENTION: A Soybean Peroxidase Gene Family and an
/ TITLE OF INVENTION: ASSAY FOR DETECTING SOYBEAN PEROXIDASE ACTIVITY
/ NUMBER OF SEQUENCES: 17
/ CORRESPONDENCE ADDRESSES:
/ ADDRESS: 1125 SO. 103RD STREET
/ CITY: OMAHA
/ STATE: NE
/ COUNTRY: US
/ ZIP: 68124-1076
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION NUMBER: US/08/671,320
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: JONDLE, ROBERT J.
/ REGISTRATION NUMBER: 33,915
/ REFERENCE/DOCKET NUMBER: 1227-001
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 402-398-9000
/ TELEFAX: 402-398-9005
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 324 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-671-320-13

Query Match 36.8%; Score 591.5; DB 2; Length 324;
Best Local Similarity 39.9%; Pred. No. 4.4e-56;
Matches 123; Conservative 62; Mismatches 116; Indels 7; Gaps 5;

QY 2 QLTPTFYDNCSPVNSIVRDTIVNELRSDPRIAASILRLHFHDFCVNGCDASILLNTTS 61
Db 22 QLQGFYANSCPAEQIVLKFVHDHINAPSLAALIRMHFDFCVRGCDASVLLNSTN 81

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 18:53:17 ; Search time 5.97557 Seconds

(without alignments)
988.783 Million cell updates/sec

Title: US-09-246-451A-17

Perfect score: 1608

Sequence: 1 MQLPTFTYDNCSEVNSIVR.....PLATGQIQLNCRVNSNS 303

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 159197 seqs, 19121497 residues

Total number of hits satisfying chosen parameters: 159157

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New.*

1: /cgn2_6/prodata/1/paa/PCR_NEW_COMB.pep.*
2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	768.5	47.8	333	6	US-10-759-602-2
2	767	47.7	318	6	Sequence 2, Appli
3	746.5	46.4	325	6	Sequence 45022, A
4	736.5	45.8	347	6	Sequence 45822, A
5	702	43.7	314	6	Sequence 40690, A
6	654	40.7	337	6	Sequence 47189, A
7	625	38.9	357	6	Sequence 39724, A
8	611	38.0	377	6	Sequence 46170, A
9	577.5	35.9	315	6	Sequence 46411, A
10	544.5	33.9	338	6	Sequence 44615, A
11	519.5	32.3	338	6	Sequence 43455, A
12	503.5	31.3	209	6	Sequence 44200, A
13	490.5	30.5	341	6	Sequence 42477, A
14	489	30.4	251	6	Sequence 44634, A
15	468	29.1	347	6	Sequence 35185, A
16	457	28.4	263	6	Sequence 46344, A
17	452	28.1	401	6	Sequence 51301, A
18	447.5	27.8	230	6	Sequence 40445, A
19	435	27.1	203	6	Sequence 39600, A
20	421	26.2	200	6	Sequence 57130, A
21	405	25.2	155	6	Sequence 55854, A
22	391	24.3	210	6	Sequence 40682, A
23	383	23.8	167	6	Sequence 36549, A
24	378	23.5	354	6	Sequence 57707, A
25	367	22.8	204	6	Sequence 45263, A
26	366.5	22.9	169	6	Sequence 57507, A
					Sequence 40120, A

27	359.5	22.4	216	6	US-10-767-701-42579	Sequence 42579, A
28	355.5	22.1	178	6	US-10-767-701-32700	Sequence 32700, A
29	351	21.8	150	6	US-10-767-701-34317	Sequence 34317, A
30	349	21.7	341	6	US-10-767-701-45520	Sequence 45520, A
31	346.5	21.5	239	6	US-10-767-701-51756	Sequence 51756, A
32	339	21.1	153	6	US-10-767-701-61014	Sequence 61014, A
33	337.5	21.0	157	6	US-10-767-701-35363	Sequence 35363, A
34	333	20.7	213	6	US-10-767-701-45881	Sequence 45881, A
35	325.5	20.2	144	6	US-10-767-701-36220	Sequence 36220, A
36	324	20.1	144	6	US-10-767-701-57979	Sequence 57979, A
37	318.5	19.8	151	6	US-10-767-701-38693	Sequence 38693, A
38	313.5	19.5	198	6	US-10-767-701-45667	Sequence 45667, A
39	306.5	19.1	230	6	US-10-767-701-43621	Sequence 43621, A
40	305.5	19.0	186	6	US-10-767-701-51813	Sequence 51813, A
41	303	19.0	159	6	US-10-767-701-32166	Sequence 32166, A
42	296.5	18.4	147	6	US-10-767-701-61688	Sequence 61688, A
43	292	18.2	158	6	US-10-767-701-60922	Sequence 60922, A
44	290.5	18.1	138	6	US-10-767-701-38689	Sequence 38689, A
45	288.5	17.9	143	6	US-10-767-701-43202	Sequence 43202, A

ALIGNMENTS

RESULT 1

US-10-759-602-2

Sequence 2, Application US/10759602

GENERAL INFORMATION:

APPLICANT: Ain-ey, Michael
Armstrong, Katherine
Belmar, Scott
Polkerts, Otto
Hopkins, Nicole
Menke, Michael A.
Paredy, Dayakar
Petolino, Joseph P.
Smith, Kelley
Woosley, Aaron

TITLE OF INVENTION: Regulatory Sequences for Transgenic Plants

NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESS:

ADDRESSEE: DowAgroSciences LLC

STREET: 9330 Zionsville Road

CITY: Indianapolis

STATE: Indiana

COUNTRY: USA

ZIP: 46268

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10759,602

FILING DATE: 16-Jan-2004

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Kraus, Eric J

TELEPHONE: 317 337 5110

TELEFAX: 317 337 4847

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 333 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Query Match: 47.8%; Score 768.5; DB 6;
Best Local Similarity 49.7%; Pred. No. 2.7e-68;
Matches 152; Conservative 52; Mismatches 99; Indels 3; Gaps 2;

NUMBER OF SEQ ID NOS: 63128

SEQ ID NO 46170

LENGTH: 357

TYPE: PRT

ORGANISM: Sorghum bicolor

FEATURE:

OTHER INFORMATION: Clone ID: SORBI-28MAY03-C17434_1.pep

US-10-767-701-46170

Query Match 38.9%; Score 625; DB 6; Length 357;

Best Local Similarity 42.7%; Pred. No. 4.8e-54;

Matches 134; Conservative 48; Mismatches 116; Indels 16; Gaps 6;

QY 7 FYDNCSPVSNIVRDTIVNELRSDPRTVSCADLLTIAAQSVTLAGGDSWVPLGRD 126

DB 29 FYNKTPSAESIVQCTVAAAGNSGVAPAIRMHFDCFVRGCDGSLIDSTANTAEK 88

QY 67 DAFGNANSARGFPVIDRMKAAYESACPRTVSCADLLTIAAQSVTLAGGDSWVPLGRD 126

DB 89 DSPANNFSLRFDDVDRAKASLEAQCFTWSCADLLAFARDVSUVTGELGYVPSGRD 148

QY 127 SLQAFLLDANLNPAPFTLPOLKDSFRNVLNRSSDLVALSGGHTFGKQCRF----- 180

DB 149 GRVSNATQATNLPFPFFFNATQVDRFASKNLT-LEDVWLGAHTLGSCHSGFAGPAN 207

QY 181 IMDELYNFSNT--GLEPDLTNTVLOT-RGLCPNLNL---SALVDFDLRTFTFDNKKY 235

DB 208 LGDELYNFSADGI-DPALS KAVAFLLKSCPSNSSOFFNTTFMDIITPDKFDNKKY 266

QY 236 VNLEEQGLIQSDQELFSSPDATDTIPLVRSPANSTQTFNFAVEAMDRMGNIPLTGTQ 295

DB 267 VGLTNLGLFESDAALLTNAIMK--ALVDSFVRNETHKKEPAKSMVKMGKIEVLGTQ 323

QY 295 GQIRLNCVWNSNS 309

DB 324 GEIRNCEVINPAS 337

RESULT 8

US-10-767-701-46411

Sequence 46411, Application US/10767701

GENERAL INFORMATION:

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

Plants and Uses Thereof For Plant Improvement

FILE REFERENCE: 38-21(53535)B

CURRENT APPLICATION NUMBER: US/10/767,701

NUMBER OF SEQ ID NOS: 63128

SEQ ID NO 46411

LENGTH: 377

TYPE: PRT

ORGANISM: Sorghum bicolor

FEATURE:

OTHER INFORMATION: Clone ID: SORBI-28MAY03-C71286_1.pep

US-10-767-701-46411

Query Match

Best Local Similarity

38.0%; Score 611; DB 6; Length 377;

Matches 133; Conservative 55; Mismatches 103; Indels 16; Gaps 7;

QY 7 FYDNCSPVSNIVRDTIVNELRSDPRTVSCADLLTIAAQSVTLAGGDSWVPLGR 125

DB 56 FYKSCPKAESIVKEFLSSAVRQNVGLAALIRVHFDCFVQGCDSVLDTPTQPSQ 115

QY 67 DAFGNAN-SARGFPVIDRMKAAYESACPRTVSCADLLTIAAQSVTLAGGDSWVPLGR 125

DB 116 LSPPNLTLPAAFAKINDIRALEQACRVSCADLLTIAAQSVTLAGGDSWVPLGR 175

QY 126 DLSQAFLLDIA-NANLPAPFTLPOLKDSFRNVLNRSSDLVALSGGHTFGKQCRFIMDR 184

DB 176 DGLAASNAVAALALPSPSTVPTLLSFLSKINLD-VTDLVALSGHTVGAHCS----- 229

QY 185 LYNFSNGLP--DPTLNTVLOTLAGLCELGNLSALVDFDLRTFTFENKYYVLEBQK 242

DB 230 --SFSNRLFPPTQDPTLNKFFAGQLYGTCTPTDTVTNTVN-DIRTPTEFNKYYVDLLNRQ 286

QY 243 GLIQSDQELFSSPDATDTIPLVRSPANSTQTFNFAVEAMDRMGNIPLTGTQGLRLNC 302

DB 287 GLFTSDQDLLEN--ATTRPIVTKFAVDQNAEEFEQFVYVYVWGQINVLTSQCGQVRANC 343

QY 303 RVVNSNS 309

DB 344 SARNAGA 350

RESULT 9

US-10-767-701-44615

Sequence 44615, Application US/10767701

GENERAL INFORMATION:

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

Plants and Uses Thereof For Plant Improvement

FILE REFERENCE: 38-21(53535)B

CURRENT APPLICATION NUMBER: US/10/767,701

NUMBER OF SEQ ID NOS: 63128

SEQ ID NO 44615

LENGTH: 315

TYPE: PRT

ORGANISM: Sorghum bicolor

FEATURE:

NAME/KEY: unsure

LOCATION: (1)-(315)

OTHER INFORMATION: unsure at all Kaa locations

FEATURE:

OTHER INFORMATION: Clone ID: SORBI-28MAY03-C56139_1.pep

US-10-767-701-44615

Query Match

Best Local Similarity

35.9%; Score 577.5; DB 6; Length 315;

Matches 119; Conservative 41; Mismatches 87; Indels 13; Gaps 5;

QY 2 QLTPFTFYDNCSPVSNIVRDTIVNELRSDPRTVSCADLLTIAAQSVTLAGGDSWVPLGR 121

DB 39 QLSSTFYDTCSPNALSTIKSGVDAAVMQEAETGASLRMHFDCFVHCGDSVLLDTSQ 98

QY 62 ERTEKDAFGNANSARGFPVIDRMKAAYESACPRTVSCADLLTIAAQSVTLAGGDSWVPLGR 121

DB 99 ---EQSSFPNKGSLRRFDV-DSIKAQVEAVCPGVVSCADLLTIAAQSVTLAGGDSWVPLGR 155

QY 122 LGRRLSQAFLDLANANLPAPFTLPOLKDSFRNVLNRSSDLVALSGGHTFGKQCRFI 181

DB 156 LGRDSTASP-PSETTDLAPTSSLOQLLSLFKNKLD-ATMVVALSGAHTIGQAQCSNF 213

QY 182 MDRLYNFSNTGLPPTLNTVLOTLAGLCELGNLSALVDFDLRTFTFENKYYVLEBQ 241

DB 214 NDHIYN-----DTNIDRAAFATSLQANCEASSTS-LAPLDTMTPTTFDNDYTNLMSQ 265

QY 242 KGLIQSDQELFSSPDATDTI 261

DB 266 KGLHSDQELFNNGSTDSTV 285

RESULT 10

US-10-767-701-43455

Sequence 43455, Application US/10767701

GENERAL INFORMATION:

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement

; FILE REFERENCE: 38-21(53535)B

; CURRENT APPLICATION NUMBER: US/10/767,701

; CURRENT FILING DATE: 2004-01-29

; NUMBER OF SEQ ID NOS: 63128

; SEQ ID NO 43455

; LENGTH: 357

; TYPE: PRT

; ORGANISM: Sorghum bicolor

; FEATURE:

; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C35094_1.pep

US-10-767-701-43455

Query Match 33.9%; Score 544.5; DB 6; Length 357;

Best Local Similarity 46.8%; Pred. No. 4.4e-46;

Matches 127; Conservative 38; Mismatches 135; Indels 11; Gaps 6;

QY 2 QLTFTFYDNCSPNVNIVRDTIVNELRSDPRFAASILRLHFHDFCFVNGCDASILLD-NTT 60

Db 34 QLKVGFTNTTCPNRAEJVRQVWTAAPANNNGVAPGLIRLHFHDFCFVNGCDASVLLSNPA 93

QY 61 SEPTKDAFGNANSARGFPVIDRMKAAVESACPRTVSCADLLTAAQGSVTLAGPSWRV 120

Db 94 GGNTERGSRANPSLRGFDVIDAAKAAVERSCPRTVSCADIVAFARDVSNLTCKLFYQV 153

QY 121 PLGRDSLOAFDLANANLPAPFTLPOLKDSFRNVLNRSSDLVALSGGTECKNOCRF 180

Db 154 PAGRRDGRVSENEADNLIGP2STAQVILGFRKNTL-VEDEVVLSGSHSTIGRSHCAS 212

QY 181 INDRLYNFSNTGLPQZTLNTTYLQTLRGLCP-LNGNLSALVDFELTPTIFDNK 233

Db 213 FLAT-RRRLADGFIISAAQALFALCPSPGQDENTTEIDVSTPAVLDDNNYKLL 268

QY 239 EEQKGLIQSQDELFPSSDADTTPLVRSFANSTQTFNFAVEAMDRMGNITELTG 293

Db 269 PMLGLHFSDQIRN---ATLAPLANAFADTELWKQKFAAAMVVMGNDIVKGTITDEI 325

QY 299 RLNCRVVNS 309

Db 326 RLNCRVVNSS 336

RESULT 11

US-10-767-701-44200

; Sequence 44200, Application US/10767701

; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement

; FILE REFERENCE: 38-21(53535)B

; CURRENT APPLICATION NUMBER: US/10/767,701

; CURRENT FILING DATE: 2004-01-29

; NUMBER OF SEQ ID NOS: 63128

; SEQ ID NO 44200

; LENGTH: 338

; TYPE: PRT

; ORGANISM: Sorghum bicolor

; FEATURE:

; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C46704_1.pep

US-10-767-701-44200

Query Match

Best Local Similarity 39.9%; Pred. No. 1.2e-43; Length 338;

Matches 125; Conservative 41; Mismatches 128; Indels 19; Gaps 7;

QY 2 QLTFTFYDNCSPNVNIVRDTIVNELRSDPRFAASILRLHFHDFCFVNGCDASILLDNTTS 61

Db 37 QLOGACYNATCPAESLIETIVHAAVRKAGNGPGLIRLFFHDFCFVNGCDASVLLDDPTG 96

QY 62 F---RT-EKDAFGNANSARGFPVIDRMKAAVESACPRTVSCADLLTAAQGSVTLAGPS 117

Db 97 TPGNRTVEKTSOPNFPFSLRGFSVINRAKRVVERRCPGTVSCADIVAFARDARIMGGR 156

QY 118 WEVPLGRDLSLOAFDLANANLPAPFTLPOLKDSFRNVLNRSSDLVALSGGTECKNO 177

Db 157 FAMPGRDLDGRVSNASEATANLPAPSENLITQLAREFASKNLT-ADDELVTLSGAHSIGRSH 215

QY 178 CRFIMDRLYNFSNTGL---PDPTLNTTYLQTLRGLCP-LNGNLSALVDFELTPTIFDNK 233

Db 216 CS-----SFANRLYPQLDATTNLVTLAARAKCPAPGKGRVVTLDFTPLQLDNG 268

QY 234 YVNLLEQKGLIQSQDELFPSSDADTTPLVRSFANSTQTFNFAVEAMDRMGNITELTG 293

Db 269 YVSNVATHEVFGSDQAL---GDRNDCAALVALYANRKMINSQKFAAMVVMGNS-EVLITG 325

QY 294 TQGOIRLNCRVVN 306

Db 326 PPGEVRLCKNKVN 338

RESULT 12

US-10-767-701-42477

; Sequence 42477, Application US/10767701

; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement

; FILE REFERENCE: 38-21(53535)B

; CURRENT APPLICATION NUMBER: US/10/767,701

; CURRENT FILING DATE: 2004-01-29

; NUMBER OF SEQ ID NOS: 63128

; SEQ ID NO 42477

; LENGTH: 209

; TYPE: PRT

; ORGANISM: Sorghum bicolor

; FEATURE:

; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C52408_1.pep

US-10-767-701-42477

Query Match

Best Local Similarity 31.3%; Score 503.5; DB 6; Length 209;

Matches 96; Conservative 31; Mismatches 48; Indels 1; Gaps 1;

QY 3 LPTFTFYDNCSPNVNIVRDTIVNELRSDPRFAASILRLHFHDFCFVNGCDASILLDNTTSF 62

Db 31 LFPQFYDHSCKPAKEIVQSIQAQAKETRMASLVRHLHFHDFCFVNGCDASVLLDNSSI 90

QY 63 RTEKDAFGNANSARGFPVIDRMKAAVESACPRTVSCADLLTAAQGSVTLAGPSWRVPL 122

Db 91 VSEKGNPNRNSLRGFEVDQIKAALEACPGTVSCADILAAARDSTSLVGGPYWDVPL 150

QY 123 GRDLSLOAFDLANANLPAPFTLPOLKDSFRNVLNRSSDLVALSGGTECKNO 178

Db 151 GRDLSLGSAGTGGSNNDIPAPNNTLPTITKEKQGLN-VVDVVVALSGGTHIRMSRC 205

RESULT 13

US-10-767-701-44634

; Sequence 44634, Application US/10767701

; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement

; FILE REFERENCE: 38-21(53535)B

; CURRENT APPLICATION NUMBER: US/10/767,701

; CURRENT FILING DATE: 2004-01-29

; NUMBER OF SEQ ID NOS: 63128

; SEQ ID NO 44634

; LENGTH: 341

; TYPE: PRT

```

; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C7853_1.pep
US-10-767-701-44634

Query Match      30.5%; Score 490.5; DB 6; Length 341;
Best Local Similarity 37.8%; Pred. No. 9e-41;
Matches 115; Conservative 48; Mismatches 134; Indels 9; Gaps 5;

QY 2 QLTPTFYNSCPNVSVIRDTIVNELRSDPRIAASILRLHFHDFCVNGCDASILLNNTTS 61
DB 36 RLSPNYRSCREVERIVSDVIAAKGRANPSTAGTILRLFFHDFCVNGCDASIVNPLSS 95
QY 62 FTEKDAFGNANSARG--FEVIDRMAAVESACPRVSCADLLTIAAQSVTLGGPSWR 119
DB 96 TAAPERAAEINLSFGDAFADAVARAKAALAESACPGVSCADVLAAARDLVLGLGPRPP 155
QY 120 VPLGRDLSLOAFDLANANLPAFFFTLPQLKDSFRVYGLNRSSDIALSGGHTFGKNOCR 179
DB 156 VALGRDARSARCVEGNLPRTNMGARAVRULFAGKGLS-ADEMVALAGAHTVGFSHCA 214
QY 180 FMDRLYNFSNTGLP--DPTLNTTYLQTLRGLOP-LNGNLSALVCFDLRPTTFIDNKYYV 236
DB 215 EFAHRIYGVGAGAGHDPLNPEFARALQRSAGYRDTPTWISFENDIVTPSEFDEAYK 274
QY 237 NLEBKGLLQSQDELFSPPDATDTIPLVRSFANSTQTFNFAVEMDRMGN-TPLTGTOG 296
DB 275 NLZPHGLGLASDAAIWEYP---PTRVFAGRYGANSKRTAFEDFAAAMQRLGAVGKTRGQ 331
QY 297 QIRLNC 302
DB 332 VVERRC 337

RESULT 14
US-10-767-701-35185
; Sequence 35185, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 35185
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C61249_1.pep
US-10-767-701-35185

Query Match      30.4%; Score 489; DB 6; Length 251;
Best Local Similarity 44.5%; Pred. No. 8.2e-41;
Matches 118; Conservative 29; Mismatches 104; Indels 14; Gaps 6;

QY 43 HDCFVNGCDASILLNNTTSFRTEKDAFGNANSARGPPVIDRMAAVESACPRVSCADLL 102
DB 1 HDCFVNGCDASILLNNTSDTGNTEKTAQAN-KSLRGFEVIDRIKEVLIESQCPGVSCADVL 59
QY 103 TTAAGQSVTLGGPSWRVPLGRDLSLOAFDLANANLPAFFFTLPQLKDSFRVYGLNRSS 162
DB 60 ALAARDAVLLARGPYGVPLGRDGRSVDSDTFTLPPFFNTSLIKLFSGHGFT-VQ 118
QY 163 DLVALSGGHTFGKNOCRFIMDRLYNFSNTGLPDPILNTTYLQTLRGCLPELNGLSALVDF 222
DB 119 DLVALSGGHTGLIAHCG-----NFKARLAETDLDALGSSLGATCAANGDDGA-APP 170
QY 223 DLRTPTIDNKYYVNLSEQKGLIQSQDELFSPPDATDTIPLVRSFANSTQTFNFAVEM 282
```

```

DB 171 D-ETSTRFDTVYVYRELQMRRLSSDQTLFESP----ETKGIVNFMAMQAYFFVAFQCGM 226
QY 283 DRMGNTPLTGTCGQIRLNCRVNS 307
DB 227 LKMGQDLLEKGEDEGEIRHCTGVNS 251

RESULT 15
US-10-767-701-46344
; Sequence 46344, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 46344
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C9288_1.pep
US-10-767-701-46344

Query Match      29.1%; Score 468; DB 6; Length 347;
Best Local Similarity 35.8%; Pred. No. 1.6e-38;
Matches 116; Conservative 53; Mismatches 107; Indels 50; Gaps 9;

QY 2 QLTPTFYNSCPNVSVIRDTIVNELRSDPRIAASILRLHFHDFCVNGCDASILLNNTTS 61
DB 45 ELSVYFHVSCPLETIVSAVDAAEQNVRLTAGLRLVFFHDFCFQGCCDASILLD---- 100
QY 62 FTEKDAFGNANSARGFP-----VIDRMAAVESACPRVSCADLLTIAAQSV 110
DB 101 -----NGERGLPPNWGLQBEAVQIVEDIRAKVHAACGPTVSCADITVLATRDV 149
QY 111 TLAGGSPMRVPLGRDLSLOAFDLANAN---LPAPFFTLPOLKDSFRVYGLNRSSDLVA 166
DB 150 SLGGSPFTVPLGRDLSVAP---ASSNDVFTLPPTSTVDALLSAFASKNLSDDPDLVA 205
QY 167 LSGGHTFGKNOCRFIMDRLYNFSNTGLPDPILNTTYLQTLRGCLPELNGLSALVDFLRT 226
DB 206 LSGAHTVGRKCS-----SFGDVAGP---ATDDITRCVTATCSAAGAGDTLRDLFLT 255
QY 227 PTIFDNKYV--NLEBKGL-LIQSDQELFSPPDATDTIPLVRSFANSTQTFNFAVEM 283
DB 256 PAVFDNLYFIETLTKXKNGVMLPSDQGLATDP---RTSLWLVQGFADNHNWFFDQFGTSMV 312
QY 284 RMGNITPLTGTCGQIRLNCRVNS 309
DB 313 KMSQLKGPQNVGEIRNCLRPNTNS 338

Search completed: April 6, 2004, 19:11:36
Job time : 7.97557 secs
```